MITIGATING PUBLIC SPEAKING ANXIETY USING VIRTUAL REALITY AND POPULATION-SPECIFIC MODELS

A Thesis

by

MEGHA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Theodora Chaspari
Co-Chair of Committee,	Tracy Hammond
Committee Member,	Amir Behzadan
Head of Department,	Dilma Da Silva

August 2019

Major Subject: Computer Science

ABSTRACT

In the education and workplace landscape of the 21st century, it is often said that a person is only as valuable as the ideas s/he has and can share. Public speaking skills are essential to help people effectively exchange ideas, persuade, inform their audiences as well as make a tangible impact. They also plays a vital role in one's academic and professional success. However, research shows that public speaking anxiety (PSA) ranks as a top social phobia among many people and tends to be aggravated in minorities, first generation students, and non-native speakers. This research aims at mitigating this anxiety by utilizing physiological (cardiovascular activity, electrodermal activity etc.) and acoustic (pitch, intonation, etc.) indices captured from wearable devices and virtual reality (VR) interfaces to quantify and predict PSA. This work also examines the significance of individual-specific factors, such as general trait anxiety and personality metrics, as well as contextual factors, such as age, gender, highest education, and native language, receny of public speaking in moderating the association between bio-behavioural (physiological and acoustic) indices and PSA. The individual-specific information is used to develop population-specific machine learning models of PSA. Results of this research highlight the importance of including such factors for detecting PSA with the proposed population-based PSA models yielding Spearman's correlation of $0.55\ n(p < 0.05)$ between the actual and predicted state-based scores. This work further analyzes whether systematic exposure to public speaking tasks in a VR environment can help alleviate PSA. Results indicate that systematic exposure to public speaking in VR can alleviate PSA in terms of both self-reported $(p < 0.05)$ and physiological $(p < 0.05)$ indices. Findings of this study will enable researchers to better understand antedecedents and causes of PSA as well as lay the foundation toward developing adaptive behavioural interventions for social communication disorders using systematic exposure (e.g., through VR stimuli), relaxation feedback, and cognitive restructuring.

DEDICATION

To my mother, Sangita and father, Anil.

ACKNOWLEDGEMENTS

I would like to begin by expressing my heartfelt gratitude to my advisor Dr. Theodora Chaspari who has helped me tremendously throughout this research journey. Dr. Chaspari has been the most patient, helpful and inspiring mentor. I am forever grateful that she decided to be my mentor and made me a better researcher. I would like to thank Dr. Amir Behzadan for his constant support and guidance in this research, his suggestions played a key role in improving this work. I would also like to thank Dr. Tracy Hammond for serving on my committee and for her constant support.

I would like to thank my fellow researchers at HUBBS and CIBER lab: Kexin Feng, Krit Gupta, Niraj Goel, Projna Promita, Md Sakib Nazmus, Nipun Debnath, and Jinwoo Kim for their constant support and encouragement. I also want to thank my friends, Vaishnavi, Pranitha, Sanjana, Priya and Anjali for being my emotional support pillars throughout this journey.

Finally, I am thankful for my parents who have always taught me the value of education and hard work. It is due to their sacrifices, undying love, and support that I have succeeded in achieving this goal.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

The thesis committee for this work include Professor Theodora Chaspari (Chair) and Professor Tracy Hammond (Co-Chair) of the Department of Computer Science at Texas A&M University and Professor Amir H. Behzadan (Member) of the Department of Construction Science.

The user study for this research was conducted with the help of Md Sakib Nazmus,Kexin Feng, Krit Gupta, and Projna Paromita, Professor Theodora Chaspari and Professor Amir H. Behzadan. Md Sakib Nazmus also took part in pre-processing the raw dataset used in this research. All other work conducted as part of this thesis was completed by the author independently.

This work has been supported by the Engineering information Foundation (EiF) through grant number (18.02). The author gratefully acknowledges the support from EiF. Any opinions, findings, conclusions, and recommendations expressed in this thesis are those of the author and do not necessarily represent those of the EiF.

Funding Sources

Graduate study for the author was supported by a research assistantship from Texas A&M University .

NOMENCLATURE

PSA	Public Speaking Anxiety
VR	Virtual Reality
ANS	Autonomic Nervous System
EDA	Electrodermal Activity
GSR	Galvanic Skin Response
HRV	Heart Rate Variability
BVP	Blood Volume Pulse
PPG	Photoplethysmogram
BSA	Behavioral Speech Anxiety
ECG	Electrocardiogram
CWP	Chest Worn Physiological
WWP	Wrist Worn Physiological
PSim	Presentation Simulator
VAD	Voice Activity Detection
SCL	Skin Conductance Level
SCR	Skin Conductance Response
IBI	Inter Beat Interval
HR	Heart Rate
FNN	Feedforward Neural Network

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iii

ACKNOWLEDGEMENTS iv

CONTRIBUTORS AND FUNDING SOURCES v

NOMENCLATURE vi

TABLE OF CONTENTS vii

LIST OF FIGURES x

LIST OF TABLES xii

1. INTRODUCTION 1

1.1 Public speaking anxiety 1
1.2 Public speaking anxiety background 2
1.2.1 State-trait model of public speaking anxiety 2
1.2.2 3 systems model of public speaking anxiety 2
1.3 Prior work 4
1.3.1 Capturing public speaking anxiety 4
1.3.2 Individual variability in public speaking anxiety 5
1.3.3 Systematic exposure to public speaking anxiety 6
1.4 Research objectives and contributions of this research 7
1.4.1 Research aims 7
1.4.2 Proposed approach 7
1.4.3 Expected contributions of this research 8

2. EXPERIMENTAL PROCEDURE 10

2.1 User study structure 10
2.1.1 Presentation tasks 10
2.1.2 PRE & POST treatments 11
2.1.3 TEST treatments 12
2.2 Wearable devices 13
2.2.1 Wrist-worn physiological (WWP) measures from empatica e4 13
2.2.2 Chest-worn physiological (CWP) measures from actiwave cardio monitor ... 14

2.2.3 Acoustic measures from microphone device 14
2.2.4 Oculus rift headset & presentation simulator 15
2.3 Self-assessment questionnaires 16

3. METHODOLOGY 24

3.1 Data pre-processing 24
3.2 Bio-behavioral measures 25
3.2.1 Empatica e4 measures 25
3.2.2 Actiwave cardio monitor measures 27
3.2.3 Acoustic measures 29
3.3 Individual and contextual factors 29
3.4 Estimation of public speaking anxiety from bio-behavioral indices 30
3.4.0.1 Correlation analysis 30
3.4.0.2 Regression analysis 31
3.5 Effect of individual-specific factors on public speaking anxiety 31
3.5.0.1 Linear regression with interaction effects 31
3.5.1 Estimation of public speaking anxiety from bio-behavioral measures augmented with individual and contextual factors 32
3.5.2 Group-specific clustering 32
3.5.3 Identifying public speaking anxiety differences between groups of participants 32
3.5.4 Group-specific public speaking anxiety models 33
3.6 Examining effect of VR stimuli on PSA 34
3.6.1 Comparing PSA before and after the VR sessions 34
3.6.2 Effect of VR environment on bio-behavioral signals 34
3.6.3 Identifying how PSA is affected by various VR settings 35
3.6.4 5-month follow-up 35

4. RESULTS 36
4.0.1 Estimation of PSA from bio-behavioral indices 36
4.0.1.1 Correlation analysis 36
4.0.1.2 Regression analysis 36
4.0.2 Effect of individual-specific factors on PSA 38
4.0.2.1 Linear regression with interaction effects 38
4.0.2.2 Estimation of PSA from bio-behavioral measures augmented with individual and contextual factors 39
4.0.2.3 Group-specific clustering 40
4.0.2.4 Identifying PSA differences between groups of participants 41
4.0.2.5 Group-specific PSA models 42
4.0.3 Examining effect of VR stimuli on PSA 43
4.0.3.1 Comparing PSA before and after the VR sessions 43
4.0.3.2 Effect of VR environment on bio-behavioral signals 45
4.0.3.3 Identifying how PSA is affected by various VR settings 46
4.0.3.4 5-month follow-up 47

5. DISCUSSION 49

6. CONCLUSIONS AND FUTURE WORK 51

6.1 Conclusions 51
6.2 Directions for future work 51

REFERENCES 53

LIST OF FIGURES

FIGURE | Page

2.1 TEST treatment in session with the participant using the Oculus headset and the Presentation simulator application to conduct a presentation in an immersive virtual environment. 12

2.2 (a) Empatica e4 wristband used during PRE, POST and TEST treatments to capture the participants electrodermal activity (EDA) signals 13

2.3 (a) Actiwave cardio monitor and (b) Audacity software used during PRE, POST and TEST treatments to capture the participants heart activity and speech signals respectively. 15

2.4 Spearman correlation heat maps for self- assessment for both PRE and POST treatments 23

3.1 Example of dropout removal and low-pass filtering ($M = 8$) for the electrodermal activity (EDA) signal using varying neighborhood window lengths (N). 25

3.2 Example of various physiological activity signals captured during PRE sessions via the empatica e4, as displayed within the e4 data analysis interface. 26

3.3 Group-specific public speaking anxiety (PSA) models, implemented through feed-forward neural network (FNN) fine-tuning. A general FNN, trained on all participants, is adapted for each group, as defined by individual and contextual factors. 33

4.1 (a) Histograms of self-reported Communication Anxiety Inventory (CAI) scores gathered during PRE session. (b) Histograms of mean heart rate measures captured via chest worn wearable device during PRE session. 37

4.2 Example of electrodermal activity (EDA) during TEST treatments 1 to 4. 37

4.3 Decision trees created from physiological features extracted from the separate wearable modalities (a) CWP and (b) WWP as well as their combination (c) CWP & WWP. 38

4.4 Predicted values of PRE session's state-based anxiety using the interaction based linear regression model for low/high score of trait-based Communication Anxiety Inventory (CAI) and physiology (RMSSD) measured via chest worn device. 40

4.5 (a) Groups of participants as determined by K-Means clustering performed on the first two principal component dimensions of the individual factors. (b)-(c) Pairwise plots of individual factors: trait anxiety and Big Five Inventory (BFI) personality metric agreeableness/extraversion with grouping governed by K-means clustering on principal component dimensions of all the individual factors. 41

LIST OF TABLES

TABLE Page

2.1 Data collection settings 11

4.1 Spearman's correlation between actual and estimated state-based anxiety using wrist-worn physiological (WWP), chest-worn physiological (CWP), and acoustic measures. 39

4.2 Spearman's correlation between actual and predicted state-based anxiety based on linear regression per modality and their combination with individual/contextual factors. 42

4.3 ANOVA & T-test results for measuring significant differences in public speaking anxiety (PSA) between various groups of individuals with respect to self-reports and bio-behavioural indices. 43

4.4 Spearman's correlation between the actual and predicted state-based anxiety measures based on the group-specific feed-forward neural network (FNN) models. 44

4.5 T-test results comparing significant differences between PRE and POST, before and after the virtual reality (VR) sessions, with respect to self-reported and bio-behavioral measures. 45

4.6 Pearson's correlation between bio-behavioral measures and state-based anxiety scores performance scores and virtual reality (VR) immersiveness/sense scores during TEST sessions. 46

4.7 Mean of state-based anxiety scores in the TEST sessions for all participants grouped based on 12 virtual reality (VR) environments provided by presentation simulator software. 47

4.8 Comparison of various state-based anxiety scores across participants between PRE treatments and 5-month follow-up survey. 48

1. INTRODUCTION

1.1 Public speaking anxiety

In the education and workplace landscape of the 21st century, communication is key and therefore it is often said that a person is as valuable as the ideas s/he has and can successfully share. Public speaking skills play a vital role in one's academic and professional success [1]. They can affect one's everyday interactions, help them in effectively exchanging their ideas [2], persuade, motivate, inform their audiences and make a tangible impact.

However, research shows that public speaking anxiety (PSA) ranks as a top social phobia among many people [3]. Individuals with PSA earn 10% fewer wages, are 10% fewer likely to graduate college, and have 15% fewer chance to obtain a leadership position compared to their peers [1]. In addition to this, according to a 2014 Forbes survey, 70% of employees who give presentations agree that public speaking skills are crucial to their success at work [4]. More recent statistics show that 57% of men and 65% of women in the U.S. view public speaking as one of their most common fears, indicating the actuality of this challenge [5]. PSA further tends to be aggravated for minorities, first-generation students, and non-native speakers [6–8]. A major cause of anxiety during public speaking is related to the novelty and uncertainty of the task, which can be alleviated through the exposure to public speaking experiences and gradual change of the negative perception related to this situation [9].

Over time, PSA has been identified via multiple names: stage fright, speech anxiety, audience anxiety, and performance anxiety. However in a more general sense, PSA is a specific type of communication-based disorder where individuals experience increased physiological arousal (e.g., heart rate, sweat activity) , negative cognition/thoughts (e.g., I am performing poorly) and/or tangible behavioral reactions (e.g., trembling) in response to a real or anticipated enactment of oral presentation [10]. High levels of PSA can have a detrimental effect on an individual's professional and academic life with a long-lasting negative impact that can render them more avoidant towards this task [1, 11–13].

In light of these problems, the primary objective of this work is to help mitigate PSA by (i) quantifying and predicting PSA using an individual's bio-behavioral signals (physiological and acoustic signals), (ii) examine the role of individual and contextual factors in predicting an individual's PSA and (iii) identify if systematic exposure to artificial public speaking stimuli (accomplished via virtual reality (VR)) can in fact help in alleviating real-life PSA. This investigation would consequently lay the foundation for creating in-the-moment VR interventions for public speaking which would predict state-based anxiety from physiological signals and automatically provide personalized feedback to the user. This feedback, in turn, is likely to reconstruct individuals' negative perception of public speaking by teaching them skills they may not have in their repertoire and alleviate their long-term trait-based anxiety. This chapter introduces some of the key concepts of PSA as well as discusses previous work done in the sphere of PSA quantification.

1.2 Public speaking anxiety background

1.2.1 State-trait model of public speaking anxiety

One of the most useful concepts for public speaking anxiety is the distinction between "trait" and "state". Public speaking anxiety experienced in the moment in a given situation is referred to as "state" whereas public speaking anxiety experienced, in general, across various situations and time periods is referred to as "trait" [10]. This distinction allows to study public speaking anxiety in a more focused manner.

Relevance to current thesis The primary objective of this work is to predict public speaking anxiety experienced in-the-moment during oral presentations (state) and examine whether public speaking stimuli provided via VR can alleviate public speaking anxiety in the short and long run (state and trait).

1.2.2 3 systems model of public speaking anxiety

The 3 systems model helps in understanding how public speaking anxiety manifests itself in humans. This model states that people in general respond to stressful situations such as public speaking in three main forms: (i) physiological, (ii) cognitive, and (iii) behavioral [14–16].

Physiological aspect of PSA The human physiological system is made of seven main systems i.e., the central, autonomic, somatic nervous, cellular and humoral systems. All of these systems regulate the human body and its response to stress [17]. Only a small subset of these physiological measures have been used in empirical PSA research [18]. The most used physiological measures are those that are representative of the autonomic nervous system (ANS). The ANS is the part of the human nervous system which regulates various body processes such as blood pressure [19] , heart rate [20], production of body fluids (sweat, saliva, etc.) [21] etc. This system works automatically in humans, i.e., without the person's conscious effort. It has two main parts: sympathetic and parasympathetic. The sympathetic system is responsible for preparing the body to react to stresses such as threat or injury. It activates what is often termed as the fight or flight response and therefore is recognized to be "associated with fear or anxiety reactions in communication settings" [18]. The parasympathetic system, on the other hand, is the part that controls functions of the body at rest. It helps in counterbalancing the sympathetic system. The ANS activity measures are also commonly used in PSA research because in terms of measurement these are "relatively strong responses which require relativity simple devices" [20].

Motivated by the above, this work utilizes the physiological measures of heart rate and sweat activity which are representative of the ANS activity. In addition to this, this work captures the speech signals of the speaker. Speech is also an inherently rich and multifaceted signal that conveys valuable information regarding a person's confidence, motivation, and affective state. Increased muscle tension when stressed causes fast palpitation of the vocal folds. Previous studies have shown that acoustic patterns (e.g., voice loudness, intonation) and paralinguistic markers (e.g., fillers, sighs, speech disfluencies) are indicative of an individual's ability to convey a clear message to the audience and are related to his/her stress levels, e.g., high number of disfluencies has been linked to increased stress, low vocal variability has been associated with poor perceptual ratings of speaking performance.

Cognitive aspect of PSA The cognitive aspect refers to the data/information collected directly from the individual performing the public speaking task. This information can be obtained by either interviews, self-report, and/or self-monitoring [22]. Most of the studies exploring PSA have relied on the self-reporting aspect of PSA [23]. This work also follows the same methodology and uses direct self-reported scores as ground truths of state-based anxiety.

Behavioral aspect of PSA Behavioral speech anxiety (BSA) is defined as “the degree of assumed speaker anxiety perceived by observers on the basis of manifest speaker behavior" [24]. It is reported that when the audience members detect BSA, the speaker’s credibility and potential speech impact suffers [18]. Mostly BSA observations (e.g., trembling) remain underutilized in PSA research [25]. This is because (i) BSA can sometimes be confused/mixed with speech quality which removes the focus from the main concentration i.e., presence and detection of PSA and more importantly (ii) measuring BSA brings with it the issue of who the judge i.e., would it be the speaker’s audience or trained third party observers [10]. Nevertheless, this work does capture the BSA measure with the help of a mix of standardized surveys and audience evaluations.

Relevance to current thesis This thesis will examine the physiological and cognitive aspect of the PSA. Physiology is quantified through bio-behavioral signals related to physiological and vocal reactivity (e.g., heart rate, sweat activity, speech intonation). Cognition is quantified based on individuals’ self-reported indices about their perception on public speaking and communication in general.

1.3 Prior work

1.3.1 Capturing public speaking anxiety

Recent developments in wearable devices have created a unique opportunity to explore PSA in various naturalistic settings and situations. Previous research has employed self-reported and signal-based measures to quantify PSA. The former refers to the speaker’s own views obtained through interviews and self-assessments , while the latter includes physiological responses to the activation of the ANS (e.g., cardiovascular and electrodermal activity), speech intonation, facial expressions, and body gestures [26, 27]. While self-reported and signal-based measures are correlated, previous studies suggest that their interaction can explain PSA better than either measure alone [28, 29]. Therefore, this work focuses on both physiology and speech signals together because of their significant correlation to PSA, privacy-preserving ability, and effortless measurement.

1.3.2 Individual variability in public speaking anxiety

Each individual can experience PSA in a different way under various settings (e.g., large or small audience). Previous studies in psychology and communication indicate that the association between physiological and self-reported state measures of PSA is moderated by a variety of psychological, cognitive, and demographic factors [30–32].

Dimberg et al. found that individuals with high trait-based public speaking fear, as reflected by general thoughts of nervousness and anxiousness, depicted increased physiological reactivity [30]. Well-prepared individuals on the presentation topic further depicted to have lower physiological reactivity compared to those who have spent less time over preparation [28]. Female speakers depicted increased self-reported and physiological anxiety when speaking in public compared to their male counter-parts [33]. Kirschbaum et al. suggested the presence of two groups of individuals (low and high responders) formed based on personality characteristics [31]. Schwerdtfeger found that the variability of self-reported state anxiety measures can be better explained when incorporating measures of trait-based anxiety, nervousness, and demographics [32]. Other studies suggest that physiological reactivity during moments of anxiety is further moderated by the knowledge of the presentation topic, novelty, impact of the presentation, and reaction and attentiveness of the audience [28, 34].

These findings suggest that there exists a complex interplay between physiology, individual-based indices, and contextual factors that contribute to PSA and therefore one general model might not be able to adequately capture the large encountered variability of physiological expressions during public speaking. While the aforementioned factors have been examined separately in previous studies, their combination has not been taken into account. In addition, previous machine learning approaches [26, 27] as stated previously, have assumed homogeneous patterns of physiological reactivity under increased PSA for all individuals. In light of this, the novelty of the current work lies in the fact that it examines the combination of various individual and contextual factors and integrates this sub-population-specific information into machine learning systems for detecting PSA from physiological and acoustic measures.

1.3.3 Systematic exposure to public speaking anxiety

Previous studies in communication and psychology indicate that PSA can be reduced via systematic exposure to public speaking encounters, which can potentially lead to the desensitization of threatening stimuli [10]. Such systematic desensitization uses exposure to public speaking stimuli in order to gradually alter the participants' perceived negative association between public speaking and anxiety [35]. Preliminary studies have explored several ways to elicit PSA. These include showing pictures of social stimuli (e.g., faces) [30], instructing speech delivery to an imaginary audience [32, 36], or presenting in front of a small-size real audience [31, 37].

Despite its effectiveness, a person's ability to create vivid mental images significantly limits the performance of such techniques [38]. This limitation can be potentially addressed through immersive experiences and VR interfaces , which can expose individuals to naturalistic public speaking stimuli via multiple virtual stimuli of greater magnitude compared to real-life [34, 39, 40]. Previous studies have also found that practicing public speaking using VR proves to be more effective in reducing an individual's PSA compared to relying on other treatments such as visualization where the audience is absent or imagined [39, 40]. A variety of recent studies have explored the feasibility of VR applications for studying and quantifying public speaking skills, performance, and anxiety [34, 39, 41–45].

VR offers an immersive experience of presenting in various public speaking stimuli without the risk of public embarrassment [46, 47]. Previous studies suggest the ability of VR interfaces to mimic threatening stimuli in a way comparable to in-vivo cues [31, 32]. Also because of its immersiveness, VR can simulate types of public speaking difficult to replicate in real-life [41, 42, 46–48]. Indicatively, Pertaub et al. [47], found that individuals experience significantly high anxiety during the exposure to negative VR audiences. North et al. [46], reasoned that VR can help individuals who have difficulty imagining public speaking scenarios. Harris et al. [42], reported that a set of four VR sessions can reduce PSA.

While previous studies in life sciences have measured PSA through self-reported and physiological indices, this work assesses the effectiveness of VR through a multi-modal set of bio-behavioral indices related to speech and physiology. Previous studies in Affective Computing have used visual and haptic feedback in order to improve public speaking skills. In Cicero, Chollet et al. [49, 50], proposed a 2D avatar augmented with visual stimuli, as provided through a color-coded visual feedback or through the interaction with the virtual audience. In the same study, public speaking performance was quantified through a set of multimodal indices related to speech, vision, and physiology. In Presentation Trainer, Schneider et al. [1], did not use an audience, but provided feedback to the user through his/her mirrored image combined with visual and haptic stimuli.

1.4 Research objectives and contributions of this research

1.4.1 Research aims

This thesis attempts to answer the following main research questions:

- Can PSA be quantified from wearable-based bio-behavioral indices?
- How to develop group-specific models of PSA?
- Does systematic exposure to public speaking encounter using VR can alleviate PSA?

1.4.2 Proposed approach

The approach of this work aims to utilize wearable technologies and VR to expose individuals to PSA stimuli and quantify and predict their PSA levels via population-specific machine learning models. To this end, physiological measures of electrodermal activity (EDA), blood volume pulse (BVP), electrocardiogram (ECG), body temperature, body acceleration, and speech are collected during public speaking presentations. The bio-behavioral indices from these signals are studied in association to retrospective self-reported state-based PSA. The proposed sub-population-specific machine learning models leverage the common information across participants and fine-tune their decisions based on specific individual (e.g., demographics) and contextual (e.g., frequency of engaging in public speaking, degree of preparation) factors. Leveraging this information stratifies groups of people with similar physiological expressions of PSA and prediction decisions are made for clusters of people with common individual-specific factors which ultimately benefits the overall system accuracy. For examining the effect of systematic exposure of VR-based public speaking stimuli on PSA, the current work compares the participants PSA before and after VR stimuli, both in terms of self-reports and bio-behavioral indices. In addition to this, the effect of the VR environment on the individuals' bio-behavioral indices and how different VR settings affect the individuals' PSA is also examined.

1.4.3 Expected contributions of this research

The main contributions of this research to the body of knowledge lies in the following: (1) Studies in Affective Computing focus on public speaking performed in front of a 2D audience in Cicero [26, 49, 50], or no audience in Presentation Trainer [1], therefore potentially lacking in terms of the user immersion, which can be provided by the VR. This research aims to harness this immersive power of VR to create realistic public speaking scenarios in order to provide naturalistic public speaking stimuli 2) Previous studies have not considered the various individual and contextual factors to quantify PSA. This research integrates these factors into group-specific machine learning models that can more accurately estimate PSA compared to general models.

The work of this research would provide the foundation for designing in-the-moment real-time feedback intervention systems for PSA. Such a cost-effective and accessible system could help minority and underrepresented students, for example, Women in Science, Technology, Engineering and Mathematics (STEM), who might exhibit an aggravated public speaking fear and alleviate their long-term PSA. In addition, this work is a stepping stone in the development of computational models of human behavior that integrate information from human perception (e.g., self-assessments), individual traits (e.g., personality), physiology (e.g., EDA) and contextual factors (e.g., environment) which would ultimately contribute in creating human-sensory integrated assistive technologies within the fields of health and education.

2. EXPERIMENTAL PROCEDURE

The data set for this research comes from a user-study which was conducted over a period of 5 months, which will be explained in this section. Appropriate IRB approval was obtained and communicated to the participants. In section 2.1, the overall structure and motivation of the user study is explained. Subsections of 2.1 explain the various session divisions within the user study. Section 2.2 lists and details the wearable devices used to capture the various bio-behavioural signals. Finally, section 2.3 explains how the different self-report assessments were used to capture the ground truth state-based anxiety levels of the speakers

2.1 User study structure

The overall aim of this user study was to have participants perform public speaking presentation in front of both real-life and virtual audiences. This would allow to assess the effect of both real-life and virtual stimuli on the participant's PSA. Participant recruitment was performed through university-level emails and advertisement. Initially 50 undergraduate and graduate-level college students were recruited for the study from Texas A&M University, aged between 18-30 years, with an equal gender distribution. For each participant the study lasted a period of 4 days. In order to increase the likelihood of observing long-term effects, each participant performed 10 separate presentation sessions during the three parts of this study: PRE, TEST and POST. The PRE and POST treatments involved a real-life audience and the TEST treatments involved virtual audiences. Due to the lengthy duration of the procedures, some participants withdrew at various stages of the study. (Table 2.1) shows the overall data collection settings and the participant statistics. In total, this user study resulted in 10,800 minutes of acoustic and physiological data from 82 real and 216 VR presentations.

2.1.1 Presentation tasks

Each of the public speaking presentation tasks under the three treatments of PRE, TEST and POST comprised of the following three phases:

Table 2.1: Data collection settings

	PRE	TEST	POST
Audience	Real	Virtual	Real
# Sessions	1	8	1
# Participants	55	38	29
# Female	23	16	13
Average age	21	21	21

- Relaxation Phase: Under this phase, participants watched a soothing video of images from nature for 5 minutes. This task was done to obtain a baseline physiological response of the participant.
- Preparation Phase: Under this phase, participants were provided with a randomly assigned news article from various topics of general interest (i.e., history, business, well-being/healthcare, entertainment/culture, technology/science, travel/nature) and were instructed that they are given 10 minutes to prepare.
- Presentation Phase: Under this phase participants presented their prepared oral presentation in front of either a real or virtual audience (real for PRE & POST ; virtual for TEST) for up to 5 minutes.

2.1.2 PRE & POST treatments

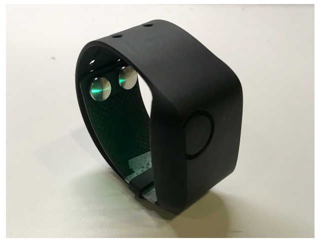
Each of the PRE and POST treatments lasted one session, each lasting an average of one hour, during which participants had to present the prepared oral presentation in front of a real-life audience in order to assess pre- and post- differences. During each of these sessions the participant went through the three stages of relaxation, preparation and presentation as described previously. The audience in these treatments comprised of professors and graduate students (on average five people) who were advised to keep a neutral demeanour throughout the presentation duration.

2.1.3 TEST treatments

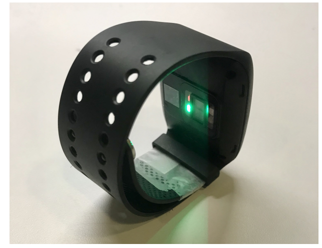
The TEST treatments comprised of eight different sessions, distributed across two days, with sessions 1-4 completed on day 1 and sessions 5-8 completed on day 2. Each of these sets of sessions took 2 hours on an average to complete. These TEST treatments were conducted between the PRE and the POST treatments, so as to enable a comparison of individual participant differences before and after the VR sessions. During each of these sessions the participants again went through the three stages of relaxation, preparation and presentation as described previously. In terms of the virtual environment, each participant was randomly assigned 8 out of 12 VR settings from various room conditions (i.e., meeting room, classroom, large hotel room), audience reactions (i.e., negative, neutral, positive), and audience size (i.e., 12, 25, 54) [34, 48].

Figure 2.1: TEST treatment in session with the participant using the Oculus headset and the Presentation simulator application to conduct a presentation in an immersive virtual environment.

(a) Placement of Empatica E4



(b) Empatica E4 front view



(C) Empatica E4 back view

Figure 2.2: (a) Empatica e4 wristband used during PRE, POST and TEST treatments to capture the participants electrodermal activity (EDA) signals

2.2 Wearable devices

2.2.1 Wrist-worn physiological (WWP) measures from empatica e4

During all the different treatments in this user study participants wore the wrist-mounted Empatica E4 [51]. The E4 wristband is a wearable research watch that offers real-time physiological data acquisition. This watch is equipped with 4 main sensors to enable physiological data collection, as listed below:

- Photoplethysmography (PPG) sensor: The PPG sensor is used to measures BVP, at a sampling rate of 64 Hz. PPG uses high precision low-intensity infrared green light sensor to detect a person's blood flow. This BVP signal can help derive the individual's heart rate or heart fluctuations. These unobtrusive PPG technology enabled sensors help in capturing the heart rate even under physical activity unlike bulky ECG sensors.

- Electrodermal sensor: This sensor measures the Galvinc Skin Response (GSR) or Electrodermal Activity (EDA) of the skin, at a sampling rate of 4 Hz. EDA refers to electrical conductance of the skin, which generally arises when the skin receives innervating signals from the brain. The E4's GSR sensor provides a way to capture this electrical conductance by passing a minuscule amount of current between two electrodes in contact with the skin. The units of measurement for this conductance is microSiemens.

- Infrared thermophile sensor: This sensor measures the temperature at a sampling rate of

4 Hz. Thermopiles, in general measure temperature by detecting an object's infrared (IR) energy. The higher the temperature, the more IR energy is emitted. It is important to note that this sensor reads the peripheral skin temperature and not the core body temperature.

- 3-axis accelerometer: This sensor captures motion-based activity of the individual at a sampling rate of 32 Hz. The accelerometer measures the gravitational force applied to each of the three spacial dimensions i.e., x, y and z.

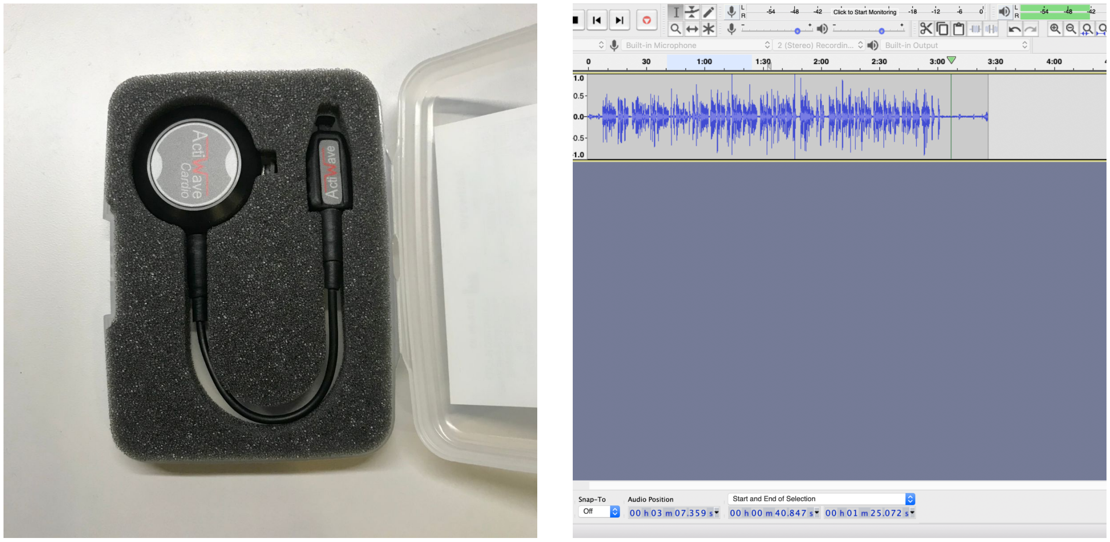
Therefore, the E4 wristband captures a number of physiological signals and achieves the goal of conducting unobtrusive real-time monitoring to obtain clinical quality physiological data. In the following discussions, the measures extracted based from the E4 wristband will be referred to as wrist-worn physiological (WWP).

2.2.2 Chest-worn physiological (CWP) measures from actiwave cardio monitor

During all the different treatments in this user study participants wore the Actiwave Cardio Monitor [52] on their chest. The Actiwave Cardio is a waterproof ultra-miniature single channel Electrocardiogram (ECG) waveform recorder with a 512 Hz sampling rate. It consists of two electrodes connected by a short lead which simply clip onto two standard ECG pads worn on the chest. It also contains a tri-axial accelerometer, the signal from which the resting body position can be determined. The small size of the device allows for continuous monitoring and unobtrusive wearability. The ECG signal captured from the actiwave is used to derive heart rate variability (HRV) measures. HRV is basically a measure of the variation in time between each heartbeat. This variation is controlled by the ANS and therefore it helps identify ANS imbalances which in turn is indicative of how stressed a person is, as discussed in detail under Section 1.2.2. In the following discussions, the measures extracted from the Actiwave cardio monitor signals will be referred to as chest-worn physiological (CWP).

2.2.3 Acoustic measures from microphone device

During all the different treatments in this user study, participants wore a creative lavalier microphone in order to capture the their live audios during their presentations. The microphone captured



(a) Actiwave cardio monitor

(b) Audacity software

Figure 2.3: (a) Actiwave cardio monitor and (b) Audacity software used during PRE, POST and TEST treatments to capture the participants heart activity and speech signals respectively.

speech signals at 16 kHz sampling rate and 16-bit encoding. The speech signals were recorded and stored via Audacity which is a free and open-source digital audio recording application software.

2.2.4 Oculus rift headset & presentation simulator

The virtual environments under the TEST treatments were created with the help of the Oculus rift headset [53] and the Presentation Simulator software [54]. The participants wore the Oculus rift headset and presented their oral presentation in front of different groups of VR audiences generated in the Presentation simulator software [54]. Oculus rift provides the VR experience by using a pair of screens that displays two images side by side, one for each eye. A set of lenses is placed on top of the panels, focusing and reshaping the picture for each eye, and creating a stereoscopic 3D image. The goggles have embedded sensors that recognize the wearer's head motions and adjust the image accordingly [53]. The leads to the user having a sensation of looking at a 3D virtual world. Presentation simulator is a VR software application designed for the Oculus rift that allows the user to present their presentations in front of a virtual audience. It contains three conference rooms representing corporate environments. The three rooms range in size from small, medium to a large conference room. The audience comprises of people with different traits and backgrounds; e.g., men, women, African American, Caucasian, young, old, casually dressed and business orientated persons. The avatars display different emotions that can be visible during the presentation which can be classified as positive, neutral and negative, which each emotion having its own sub set of specific behaviors to mimic real life situations. In addition to these two, in order to enable an immersive VR experience, the participants were also made to listen to a constant classroom-based background noise using a Youtube video [55] through the Oculus rift headset headphones.

2.3 Self-assessment questionnaires

Under all of the 3 treatments of PRE, TEST and POST , the candidates filled numerous self-assessments, both before and after their presentations. These self-assessments were used to capture the ground truths, i.e., the participants self-reported state and trait anxiety, their individual-specific and contextual factors such as age, gender, ethnicity, level of preparation etc. Following section discusses the details of the various self-assessments acquired and what each of them aimed to capture.

Self-assessment reports before the PRE and POST sessions

Participants filled the following questionnaires prior to entering the relaxation phase in both PRE and POST sessions.

- **Trait-Scale of the State Trait Anxiety Inventory (STAI)** [56] STAI is a commonly used measure of trait and state anxiety. The trait scale of the questionnaire has 20 questions for assessing trait anxiety. Some example questions are: “I worry too much over something that really doesn’t matter" and “I feel pleasant; I lack self-confidence". All items are rated on a 4-point Likert scale (e.g., from “Almost Never" to “Almost Always"). Higher scores indicate greater anxiety.

- **Trait-Scale of the Communication Anxiety Inventory (CAI)** [57] CAI also measures both trait and state anxiety. The trait scale of the questionnaire has 21 questions for assessing trait anxiety. Some example questions are: “I think I communicate effectively in one-to-one situations" and “I enjoy speaking in public; I make a good impression when I speak in public". All items are rated on a 4-point Likert scale (e.g., from “Almost Never" to “Almost Always"). Higher scores indicate greater anxiety. Pre-specified summations of certain selected items from the overall set of 21 questions provide 3 more scores: the CAI dyadic score, CAI Small group score and CAI Public speaking score.

- **Personal Report of Public Speaking Anxiety (PRPSA)** [58] focuses strictly on measuring PSA instead of broader communication apprehension. The questionnaire has 34 questions for assessing PSA. Some example questions from the assessment are: “While preparing for giving a speech, I feel tense and nervous" , “My hands tremble when I am giving a speech" and “I perspire just before starting a speech". All items are rated on a 5-point Likert scale (e.g., from “Strongly Agree" to “Strongly Disagree").

- **Brief Fear of Negative Evaluation (BFNE)** [59] PSA is also partly generated due to the perceived negative evaluation by others [60, 61]. The fear of negative evaluation consists of feelings of apprehension about others’ evaluations, distress over these negative evaluations, and the expectation that others will evaluate one negatively [62]. The big difference between fear of negative evaluation and PSA is that the former pertains to the sense of dread associated with being evaluated unfavorably while performing a public speaking task, whereas the latter refers to the affective reactions caused while performing a public speaking task. Therefore, the BFNE was employed to measure this construct of dread. The questionnaire has 12 questions. Some example questions from the assessment are: “I am unconcerned even if I know people are forming an unfavorable impression of me", “I am afraid that people will find fault with me" and “I am afraid others will not approve of me". All items are rated on a 5-point Likert scale (e.g., from “Not at all characteristic of me" to “Extremely characteristic of me").

- **Reticence Willingness to Communicate (RWTC)** [64] assess a person’s reluctance or predisposition towards communicative situations. The questionnaire has 31 questions to measure one's reluctance. Some example questions from the assessment are: “In general, I feel at ease when speaking", “I tend to postpone oral contacts as long as I can" and “Speaking in front of an audience makes me feel tense". All items are rated on a 5-point Likert scale (e.g., from “Strongly Agree" to “Strongly Disagree").

- **Demographics** is a custom made survey which captures the participant's age, biological sex, primary language, ethnicity, education, etc. This questionnaire had 11 questions. Some example questions from the assessment are: “What is the highest education level that you have completed" , “What is your primary language" and “What is your ethnicity". All items are rated on different multiple choice-based options.

- **Daily Experience questionnaire** is a custom made survey which captures the participant's daily activities which might prove to be a confounding factor in their presentation performance. This questionnaire had 7 questions. Some example questions from the assessment are: “How long ago was your last meal (including breakfast, lunch, dinner)" , “How many cups of alcoholic drinks have you consumed today" and “Has there been a significant event in the past week that could affect your performance in this task". All items are rated on different multiple choice-based options.

Self-assessment reports after the PRE and POST sessions

Participants filled the following questionnaires after finishing the presentation phase in both PRE and POST sessions.

- **State-Anxiety Enthusiasm (SAE)** captures the state-based anxiety of the participants related to the preceding public speaking encounter. The questionnaire has 20 questions to measure one's state-based anxiety. Some example questions from the assessment are: “My listeners seemed to be interested in the topic of my presentation" , “I succeeded in my task better than I had anticipated" and “I felt my hands shaking when I was speaking". All items are rated on a 5-point Likert scale (e.g., from “Strongly Agree" to “Strongly Disagree").

- **State-Scale of the State Trait Anxiety Inventory (STAI)** [56] The State scale of the STAI questionnaire has 20 questions for assessing state anxiety. Some example questions are: "I am presently worrying over possible misfortunes" and "I feel frightened; I feel upset" All items are rated on a 4-point Likert scale (e.g., from "Almost Never" to "Almost Always"). Higher scores indicate greater anxiety.

- **State-Scale of the Communication Anxiety Inventory (CAI)** [57] The state scale of the CAI questionnaire has 20 questions for assessing state anxiety. Some example questions are: "I felt tense and nervous" and "I felt self-confident while talking; I could not think clearly when I spoke". All items are rated on a 4-point Likert scale (e.g., from "Almost Never" to "Almost Always"). Higher scores indicate greater anxiety.

- **Body Sensations Questionnaire (BSQ)** [65] captures the participants physiological reactivity when involved in a public speaking task. The questionnaire has 18 questions for assessing physiological reactivity. Some example questions are: "I felt nausea" and "I was sweating; I had a dry throat". All items are rated on a 5-point Likert scale (e.g., from "Not at all" to "Extremely"). Higher scores indicate greater physiological reactivity.

- **Presentation Preparation Performance (PPP) survey** captures the participant's degree of preparation and knowledge on the topic. The questionnaire has 6 questions for assessing the level preparation. Some example questions are: "How would you rate the difficulty of the topic that was given to you to present?" and "How would you rate the level of your concentration while preparing for the presentation?". All items are rated on different multiple choice-based options.

Self-assessment questionnaires before the TEST sessions

Participants filled the following questionnaires before starting the collective TEST sessions on a particular day (collective TEST sessions : 1-4 on day 1 and 5-8 on day 2).

- **Daily Experience questionnaire** is a custom made survey which captures the participant's daily activities which might prove to be a confounding factor in their presentation performance as explained in detail under Section 2.3

- **Personal Report of Public Speaking Anxiety (PRPSA)** [58] is the same questionnaire employed after the PRE and the POST sessions.

- **Brief Fear of Negative Evaluation (BFNE)** [59] is the same questionnaire employed after the PRE and the POST sessions.

- **Big Five Inventory (BFI)** [66] BFI is a self-report inventory designed to measure a person's personality traits. The questionnaire has 44 questions that measures an individual on the Big Five Factors (dimensions) of personality (Goldberg, 1993). The Big Five personality dimensions or factors are Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness. Some example questions are: "I see myself as someone who is talkative " and "I see myself as someone who is sometimes shy, inhibited". All items are rated on a 5-point Likert scale (e.g., from "Strongly Agree" to "Strongly Disagree").

- **Memory capacity and cognitive test** An online memory capacity test and cognitive test was used to capture the participant's cognitive and memory ability since these might prove to be a confounding factor in their presentation performance.

  The memory test was a simple picture memory test where candidates were shown a series of images. If they saw an exact repeat image, they were supposed to click the image. The test provided the candidates mean reaction time and number of pictures they got correct. The cognitive test had Verbal, Numerical and Abstract subdivisions each containing 9 questions to measure the participant's cognitive aptitude.

Self-assessment questionnaires after the TEST sessions

Participants filled the following questionnaires after each of the VR TEST sessions.

- **State-Anxiety Enthusiasm (SAE)** is the same questionnaire employed after the PRE and the POST sessions.

- **VR Sense** [67] is used to identify the users' experiences of media and how present they felt under interactive virtual environments. The questionnaire has 8 questions for assessing physiological reactivity. Some example questions are: "I was distracted by the quality of the technology" and "I was aware of the real world; I found it easy to forget that I was watching a display". All items are rated on a 7-point Likert scale (e.g., from "Very aware" to "Hardly Aware").

- **Presentation Preparation Performance (PPP) survey** is the same questionnaire employed after the PRE and the POST sessions.

Participants further filled the following questionnaires after finishing the collective TEST sessions on a particular day.

- **VR Presence** [68] is used to identify how immersive was the VR for the participant and how life-like was the public speaking experience for them in the VR environment. The questionnaire has 19 questions for assessing physiological reactivity. Some example questions are: "How much were you able to control events?" and "How much did the visual aspects of the environment involve you?; How compelling was your sense of moving around inside the virtual environment?". All questions are rated on different multiple choice-based options.

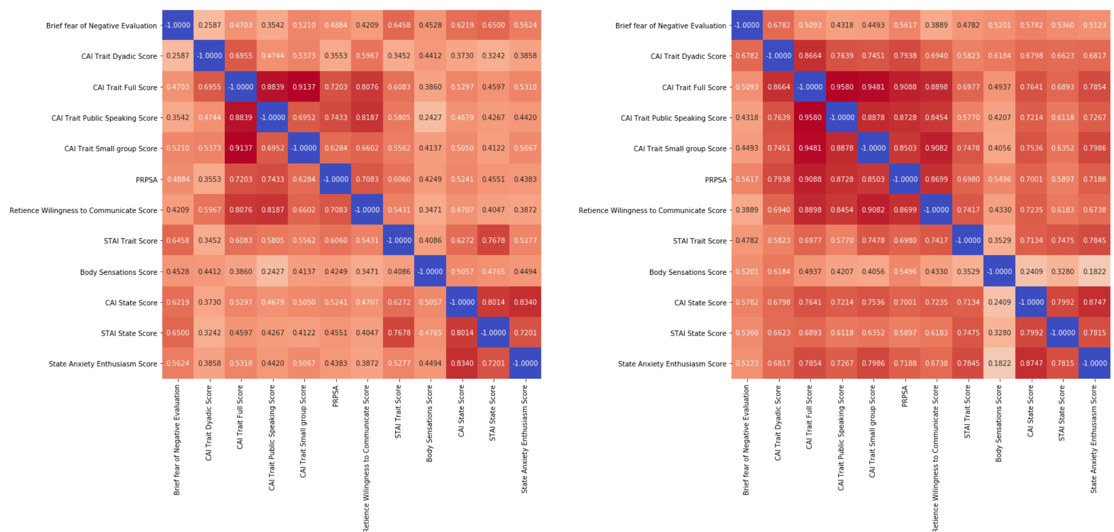
- **VR SUS Slater-Usoh-Steed (SUS) questionnaire** [69] is used to identify if participants can distinguish between real and virtual experiences. The questionnaire has 5 questions for assessing physiological reactivity. Some example questions are: "Do you have a vivid or realistic memory of the virtual environment? " and "To what extent were there times during the experience when the virtual environment was the reality for you?". All questions are rated on different multiple choice-based options.

- **State-Scale of the State Trait Anxiety Inventory (STAI)** [56] is the same questionnaire employed after the PRE and the POST sessions.

- **State-Scale of the Communication Anxiety Inventory (CAI)** [57] is the same questionnaire employed after the PRE and the POST sessions.

- **Body Sensations Questionnaire (BSQ)** [65] is the same questionnaire employed after the PRE and the POST sessions.

Within the design of the presented user study, the aforementioned self-assessments are utilized as the sole ground truths for state and trait-based anxiety. Therefore, in order to ensure that the participants provide truthful responses in the self-assessments and consequently allow one to conduct outlier detection for unreliable scores, a certain amount of redundancy was introduced in the survey questions by including similar surveys, for example, the CAI trait survey and the STAI trait survey both capture an individual's trait based anxiety and therefore have similar themed questions. Figure 2.4 showcases the Spearman correlations among the various self-assessments for both the PRE and the POST treatments. It shows high correlations between surveys which capture a participant's state and trait based anxiety, e.g., STAI trait and CAI Trait (PRE Spearman correlation $= 0.61$, POST Corr $= 0.70$) and STAI state survey and State Anxiety Enthusiasm survey (PRE Spearman correlation $= 0.72$, POST Spearman correlation $= 0.78$),etc.



(a) PRE treatment

(b) POST treatment

Figure 2.4: Spearman correlation heat maps for self- assessment for both PRE and POST treatments

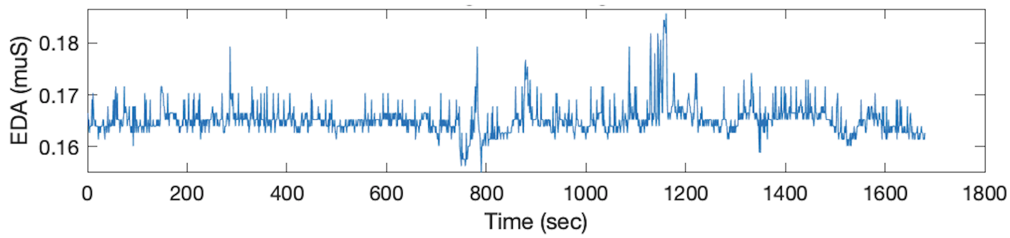
3. METHODOLOGY

This chapter describes the analysis conducted to answer the three research questions (Section 1.4.1). Section 3.1 describes the pre-processing of physiological and acoustic signals. Section 3.2 describes the various bio-behavioural features extracted, Section 3.3 outlines the various individual and contextual factors which could potentially contribute to the participant's PSA . The next few sections describe the analyses carried out to answer the three research questions of this work: (i) Can we estimate PSA from bio-behavioral indices? (Section 3.4), (ii) How do individual-specific factors contribute to PSA? (Section 3.5), and (iii) Does systematic exposure through VR alleviate PSA? (Section 3.6).

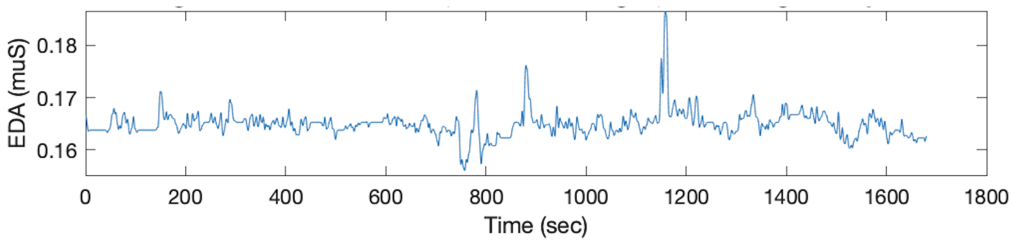
3.1 Data pre-processing

Physiological signals collected in ambulatory settings tend to depict increased levels of noise, that yield from movement artifacts, sensor misplacement, loss of electrode contact with the skin, and electrode leakage. For this reason, initially all the raw time-series physiological signals were visualized. This allowed to determine which of the captured signals depicted the expected characteristic structure and could be used for further investigation. Few examples of signals removed via visual inspection were: EDA signals which displayed constantly low (< 0.01 muS) values with no fluctuations and ECG signals which were consistently not captured through the actiwave monitor due to sensor displacement and resulted in bouts of flat ECG line.

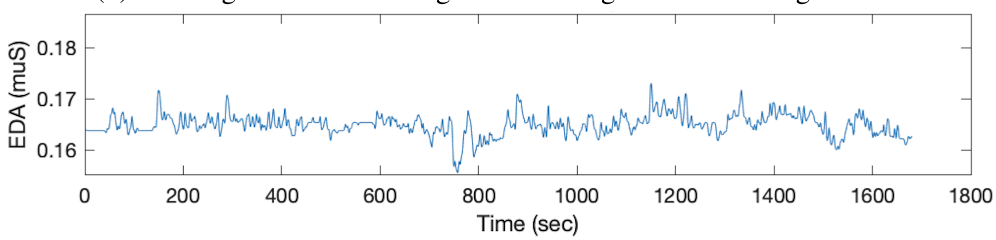
Post the visual inspection, outlier detection was performed for the EDA to detect potential dropouts. Outliers were defined as signal samples with values larger than three times the standard deviation from the median over an analysis window of 48 samples, a value visually yielding the best results (3.1). Outliers were replaced by carrying out a linear interpolation using the neighboring signal values using the Matlab *Filloutliers* function [70]. Post this, the EDA signal was treated with a Bateman low-pass filter with a 8-sample length to remove high-frequency noise. For ECG signals, high-frequency noise was removed using a low-pass finite impulse response filter of 45-



(a) Original EDA signal



(b) EDA signal after removing outliers using a $N = 24-$length window



(c) EDA signal after removing outliers using a $N = 48-$length window

Figure 3.1: Example of dropout removal and low-pass filtering ($M = 8$) for the electrodermal activity (EDA) signal using varying neighborhood window lengths (N).

samples length, followed by R-peak detection using the BioSPPy toolbox [71]. For speech signals voice activity detection (VAD) was performed to identify the presence and absence of speech using the OpenSMILE [72] toolbox.

After pre-processing, all physiological signals were segmented according to the three main task phases: relaxation, presentation preparation, and presentation.

3.2 Bio-behavioral measures

3.2.1 Empatica e4 measures

The Empatica E4 provides 4 main physiological signals i.e., EDA signal, BVP signal, 3-axis acceleration signal and temperature signal. A total of 7 features are extracted from these physiological signals as described below:

- **EDA signal metrics** : EDA or galvanic skin response (GSR) refers to the changes in an individual's sweat gland activity. The GSR signal are reflective of the intensity of one's emotional state/ emotional arousal, but not the type of emotion. Emotional arousal can be created via both positive or negative stimuli and results in an increase in the individual's skin conductance. GSR is driven autonomously by sympathetic activity which drives aspects of human behavior, as well as cognitive and emotional states [73]. Skin conductance therefore offers direct insights into autonomous emotional regulation. The EDA signal is the result of two additive processes: a tonic base level driver, which fluctuates very slowly , and a faster-varying phasic component . We consider both these tonic and phasic level components. This provides us with the following 3 EDA metrics. These metrics are extracted from the EDA signal using the Ledalab software [74].

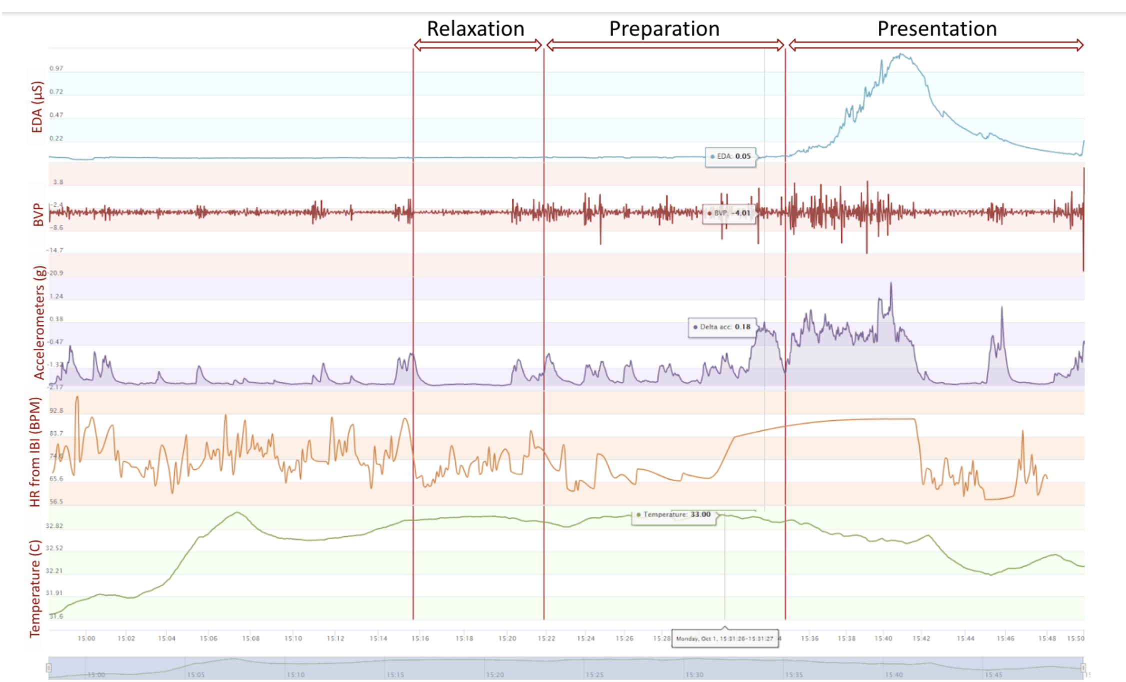


Figure 3.2: Example of various physiological activity signals captured during PRE sessions via the empatica e4, as displayed within the e4 data analysis interface.

  - mean SCL (tonic), the mean level of the EDA signal.

  - SCR frequency (phasic), the number of skin conductance responses divided by the duration of the corresponding time segment,
  - mean SCR amplitude (phasic), depicts the mean amplitude of skin conductance responses within a time segment

- **BVP signal metrics**: As stated previously, the BVP signal is obtained via the PPG technology. The BVP captures changes in blood volume in the arteries and capillaries that correspond to changes in the heart rate and blood flow.
  - Heart rate: The BVP measures heart rate based on the volume of blood that passes through the tissues in a localized area with each beat (pulse) of the heart. More precisely, heart rate is computed by detecting peaks (beats) from the BVP and computing the lengths of the intervals between adjacent beats.
  - IBI: The time between beats is measured in milliseconds (ms) and is called an "inter-beat interval (IBI)". The BVP is the input signal to the proprietary algorithm of Empatica E4 device that detects the heart beats and that provides the IBI signal as output.

- **3-axis acceleration signal metrics**: Empatica devices have an 3-axis accelerometer which measures gravitational force applied to each of the three spacial dimensions i.e., x, y and z. $l2-$norm of this 3-axis acceleration signal is used as a metric.

- **Body temperature signal metrics**: The mean temperature is considered as a metric from the peripheral skin temperature signal provided by Empatica's Infrared thermopile sensor.

3.2.2 Actiwave cardio monitor measures

The Actiwave cardio monitor provides 1 main physiological signal i.e., the ECG signal. Two types of feature sets can be extracted from this ECG signal as described below:

- ECG signal: Electrocardiography (ECG) is a graphical representation of the electrical activity of the heart over a period of time. The QRS detection is a very important step in ECG signal processing. The bio signal processing toolbox Biosspy in python was used to extract a series of successive R-peak location from this raw ECG signal. Next, the pyhrv heart rate variability toolbox [75] was used to compute the series of NN intervals from the R-peak locations.

Next, HRV features are extracted from the NN-intervals. HRV is the measure of the variation in time between each heartbeat. This variation is controlled by the ANS and therefore it helps identify ANS imbalances which in turn is indicative of how stressed a person is.

- Time domain features: The hrv-analysis toolbox [76] is used to extract the set of HRV time domain features from the NN intervals.
  - RMSSD: The square root of the mean of the sum of the squares of differences (RMSSD) between adjacent NN-intervals. The RMSSD reflects high frequency (fast or parasympathetic) influences on HRV (i.e., those influencing larger changes from one beat to the next). The RMSSD measure reflect the body's parasympathetic activity, which contributes to one's self-regulation ability [77].
- Frequency domain features: The HRV-analysis toolbox [76] is used to extract the set of HRV Frequency domain features from the NN intervals.
  - Low-frequency (LF) energy: LF reflects the variance or power in HRV in the low Frequency domain (.04 to .15 Hz). Reflects a mixture of sympathetic and parasympathetic activity, but more prominently in slightly long-term recordings, it reflects sympathetic activity.
  - High-frequency (HF) energy: HF reflects the variance or power in HRV in the High Frequency (.15 to .40 Hz by default). Reflects fast changes in beat-to-beat variability due to parasympathetic activity .
  - LF-HF ratio : While most investigators also use the LF-HF ratio as a representative of the sympathetic activity , the precise role of this measures tends to be unclear [78].

3.2.3 Acoustic measures

The microphone captures the speech signals. A total of 7 acoustic features are extracted from these speech signals using the OpenSMILE [72] toolbox. These features were computed over a 30-millisecond analysis window and were averaged over the speech segments of each audio file.

- Speech Signal:
  - Root Mean Square (RMS) energy: The RMS (Root-Mean-Square) value is the effective value of the total signal waveform. It is really the area under the curve. Therefore, in speech it is the power that is delivered.
  - Fundamental frequency (F0): It is basically the inverse of the pitch period length. It is a measure of how high or low the frequency of a person's voice sounds. Its psychologically correlated with pitch.
  - Number of pauses: This measure reflects the fluency of the speaker.
  - Zero Crossing Rate (ZCR): This measure represents the sign-change rate of speech.
  - Jitter and shimmer: Jitter and shimmer are the two common perturbation measures in acoustic analysis. Jitter is a measure of frequency instability, while shimmer is a measure of amplitude instability.
  - Voicing probability: This measure represents the probability of voice activity based on autocorrelation function.

3.3 Individual and contextual factors

A total of 14 individual and contextual factors were used to model the inherently high variability across individuals and across various conditions, as obtained from the participants' self-reports (Section 2.3). Significant differences between individuals with respect to their self-reports and bio-behavioral indices were studied based on these factors. These factors were further examined in terms of their ability to moderate the association between bio-behavioural indices and state-based PSA. Contextual factors include:

- Gender
- Age
- Native language
- Ethnicity
- Highest educational degree achieved
- Degree currently being pursued
- Majoring in which subject
- Recency of public speaking experience
- Self-reported level of preparation and knowledge on the presentation (PPP)

Individual factors include:

- Personality metrics (BFI questionnaire)
- Trait-based general anxiety levels (STAI Trait)

3.4 Estimation of public speaking anxiety from bio-behavioral indices

This section primarily aims to understand if bio-behavioral indices can be used to quantify and measure PSA.

3.4.0.1 Correlation analysis

A preliminary inspection of the self-reports and physiological measures captured via the CWP and the WWP device was carried out. In addition to this, Pearson's correlation was carried out between bio-behavioural indices and self-reported anxiety scores. Pearson's correlation is a measure of the strength of a linear association between two variables, it attempts to draw a line of best fit through the data of two variables. Pearson's correlation was used to examine the degree of association between the various bio-behavioural indices (Section 3.2) and state-based PSA scores.

These scores were obtained by the State-based CAI questionnaire (Section 2.3). This correlation analysis was done solely for the PRE and the POST treatments.

3.4.0.2 Regression analysis

Linear regression was performed to estimate/predict the individuals' self reported state-based PSA, as reported from the CAI State survey (Section 2.3). The input features for the regression models were based on each of the different wearable modalities (CWP, WWP, Audio) and their combination (Section 3.2). Each of the regression model was evaluated through a leave-one-speaker-out (LOSO) cross-validation, according to which data from one speaker were included in the test set, while data from the remaining speakers were used for training. The estimated/predicted and the actual state-based PSA values were compared using Spearman's correlation. These state-based anxiety predictions were done solely for the PRE and the POST treatments.

3.5 Effect of individual-specific factors on public speaking anxiety

This section primarily aims to understand how individual and contextual factors affect an individuals' PSA and if adding these factors in addition to bio-behavioral indices increases the predictive power of PSA models.

3.5.0.1 Linear regression with interaction effects

A linear regression model with interaction effects was used to predict participants' self-reported state-based anxiety y from bio-behavioral index x, individual or contextual factor c, and their interaction, as follows:

$$y = a_1 x + a_2 c + a_3 x \cdot c \tag{3.1}$$

In (3.1), a_1, a_2 quantify the association of state-based anxiety with the bio-behavioral index and the individual-specific moderating factor, respectively. The coefficient a_3 reflects how the individual-specific factor moderates the association between physiology/speech and state-based anxiety. If $a_3 > 0$, a stronger association exists between bio-behavioral indices and state-based anxiety for participants with higher levels of trait-based anxiety or preparation compared to their counter-parts.

3.5.1 Estimation of public speaking anxiety from bio-behavioral measures augmented with individual and contextual factors

Linear regression was conducted based on the bio-behavioral features (CWP, WWP, Audio) and their combination with the individual and contextual factors. The original 18-dimensional feature vector of bio-behavioral indices as discussed in Section (3.2) was augmented by individual (trait-based anxiety from STAI, personality scores from BFI) and contextual factors (age, gender, native language, ethnicity, recency of public speaking experience, highest education achieved, currently pursuing degree). Each factor added one feature to the final feature set. The goal of each regression was to identify whether including individual indices affects the prediction of state-based anxiety-based on the bio-behavioural measures captured from the different wearable modalities. Each regression model was evaluated through a leave-one-speaker-out (LOSO) cross-validation by computing the Spearman's correlation between the actual and estimated state-based PSA.

3.5.2 Group-specific clustering

Different individuals are likely to experience different patterns of anxiety in various settings [10]. In order to integrate such individual and contextual differences into machine learning models, participants were clustered into different groups based on their individual and contextual factors (Section 3.3). This allows to understand the subsets of factors affecting the state-based anxiety and its association with bio-behavioural signals. Principal Component Analysis (PCA) was applied on the individual and contextual factors to reduce their dimensionality. Next, K-Means clustering was performed on the first two PCA dimensions, to obtain $K = 4$ groups of participants. The value of K was empirically determined based on the number of data samples.

3.5.3 Identifying public speaking anxiety differences between groups of participants

Statistical analysis was used to identify any significant differences between groups of participants with respect to their bio-behavioral indices and self-reported scores. Grouping was performed based on the individual and contextual indices (Section 3.3). A t-test was used in the case of two groups such as gender (male, female) , while an analysis of variance (ANOVA) was conducted when more than two groups were present such as in age (18 to 22, 22 to 26, 26 to 30), native language (Hindi, English, Spanish, Other), etc.

3.5.4 Group-specific public speaking anxiety models

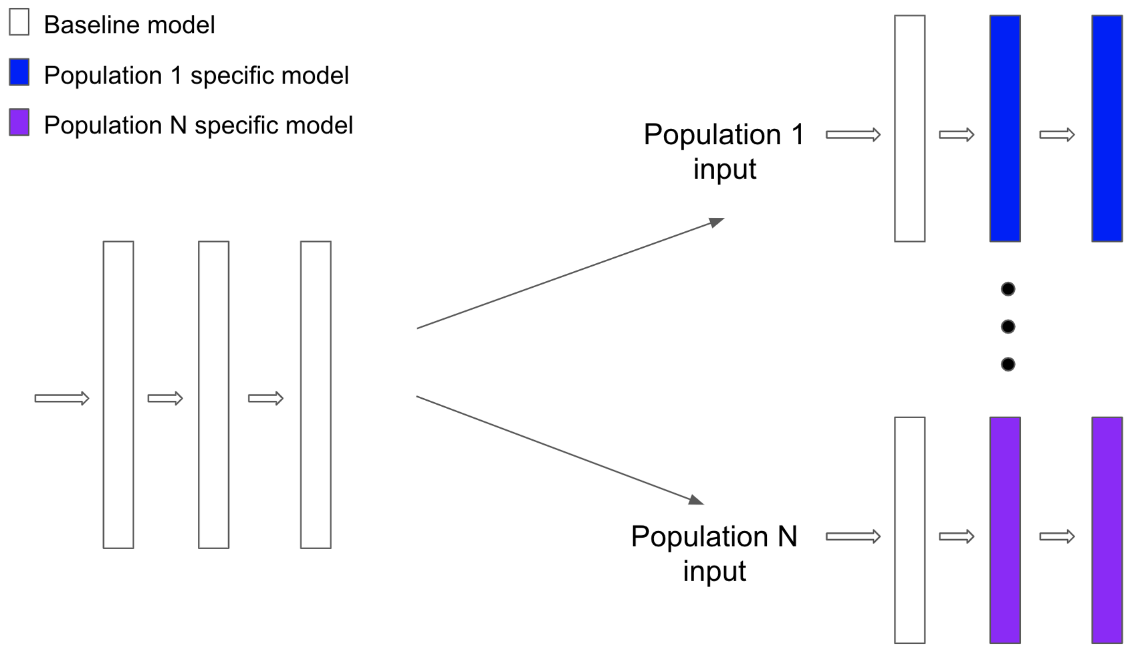


Figure 3.3: Group-specific public speaking anxiety (PSA) models, implemented through feed-forward neural network (FNN) fine-tuning. A general FNN, trained on all participants, is adapted for each group, as defined by individual and contextual factors.

A feed-forward neural network (FNN) was trained based on all data samples as a baseline model to estimate state-based PSA from individuals' bio-behavioral indices. The FNN comprised of one hidden layer and was trained with a learning rate of 0.01, providing a general PSA estimation for all participant. Next, group-specific fine-tuning was performed, based on which samples from each group of participants were used to fine-tune the hidden and output layer of the baseline FNN, resulting in group-specific PSA estimations (Fig. 3.3). The learning rate during fine-tuning was 0.001, providing fine-grain learning of the FNN parameters. Note that FNN fine-tuning is not performed for clusters with less than 3 data points, since this will not provide an adequate amount of data for re-training the last FNN layers. Three different types of group-specific PSA models were created based on three different group clustering criteria (individual, contextual, and their combination). Evaluation was performed through Spearman's correlation values using the state-based PSA estimations obtained from LOSO. Therefore, in this approach population-specific models are created based on the various individual-specific factors which possibly affect PSA to predict state-based anxiety. Furthermore, it is analyzed how these population-specific FNN models perform compared to a solely bio-behavioural-based baseline FNN models in predicting state-based anxiety.

3.6 Examining effect of VR stimuli on PSA

3.6.1 Comparing PSA before and after the VR sessions

It was examined whether frequent exposure to VR stimuli, provided by the 8 VR sessions in TEST , affected the participants' PSA. Statistical analysis through paired t-test was employed to determine significant differences between the PRE and the POST with respect to participants' trait-based and state-based self-reported anxiety, as well as their bio-behavioral measures (CWP, WWP, Audio). T-tests were also conducted with respect to each of the physiology and speech features to identify any significant effects of VR on these bio-behavioural signals of participants during public speaking.

3.6.2 Effect of VR environment on bio-behavioral signals

The notion of "immersiveness" in the VR environment was examined by analyzing how this can affect the bio-behavioral signals of the participants. Correlation analysis through Pearson's correlation was employed to determine significant correlations/associations between the various bio-behavioral measures (CWP, WWP, Audio) and the self-reported state-based anxiety scores, performance scores and VR immersiveness/sense scores.

3.6.3 Identifying how PSA is affected by various VR settings

As mentioned previously the presentation simulator software provides various VR environments/settings for the user to present in front of. The effect of these 12 settings on the self-reported state-based anxiety scores was examined. The 12 settings comprise of 4 main changing variables i.e., Room type, Audience Size, Audience reaction and the number of females in the room. Analysis of variance (ANOVA) tests were performed to to identify if these 4 different variables had any significant effect on the participants' bio-behavioural and self-reported anxiety indices.

3.6.4 5-month follow-up

In order to evaluate the long-term effects of systematic exposure on trait-based PSA via VR-based public speaking stimuli, a 5-month follow-up survey was done with the user study participants. This follow-up captured the participants' trait-based anxiety levels via the self-assessments explained previously. Statistical analysis through paired t-test was employed to determine significant differences between the participants' trait-based anxiety captured in PRE treatments and the trait-based anxiety captured in the aforementioned 5-month follow-up surveys.

4. RESULTS

This chapter discusses the results of the various analyses described in Section 3. For the analyses conducted solely on PRE and POST treatments, the results are reported separately for both the PRE and POST treatments, since the TEST treatment (consisting of 8 VR sessions) that took place in between them, renders them substantially different and indicative of PSA evolution due to VR.

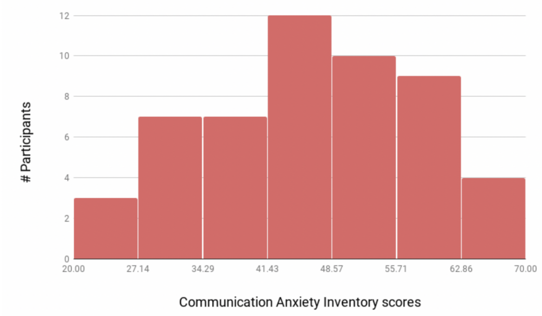
4.0.1 Estimation of PSA from bio-behavioral indices

4.0.1.1 Correlation analysis

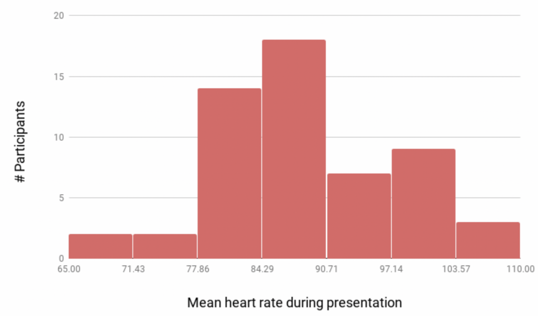
Preliminary inspection of the responses gathered from the self-reports, specifically the CAI survey, indicates that there is a substantial proportion of individuals in the collected dataset that suffer from general communication anxiety (Figure 4.1a)). Similarly, inspection of the physiological measures obtained via the CWP device indicates that there are a substantial proportion of individuals in the collected dataset that experienced increased heart activity during the presentation tasks (Figure 4.1b)). A visual inspection of physiological indices also indicates increased physiological activity for EDA during presentation tasks in PRE, TEST and POST treatments (Figure 4.2). The correlation analysis conducted via Pearson's correlation did not provide numerous significant results, however it did indicate significant associations between few important bio-behavioral indices, which are highly indicative of ANS activity, and self-reported state-based anxiety scores. For example, participants with high self-reported trait anxiety depict higher EDA during presentation (Pearson's correlation= 0.46, p= 0.013) as well as high HRV LF value (which represents the sympathetic system and thus the flight/fight response) ($r = 0.27, p = 0.05$)

4.0.1.2 Regression analysis

Linear regression results indicate significant associations between the proposed bio-behavioural indices and the self-reported state-based anxiety scores. (Table 4.1) shows the correlations between actual and estimated state-based anxiety using wrist-worn physiological (WWP), chest-worn physiological (CWP), and acoustic measures. In terms of singular modalities, CWP and Acoustic



(a) Communication Anxiety Inventory (CAI) scores



(b) Mean heart rate measures

Figure 4.1: (a) Histograms of self-reported Communication Anxiety Inventory (CAI) scores gathered during PRE session. (b) Histograms of mean heart rate measures captured via chest worn wearable device during PRE session.

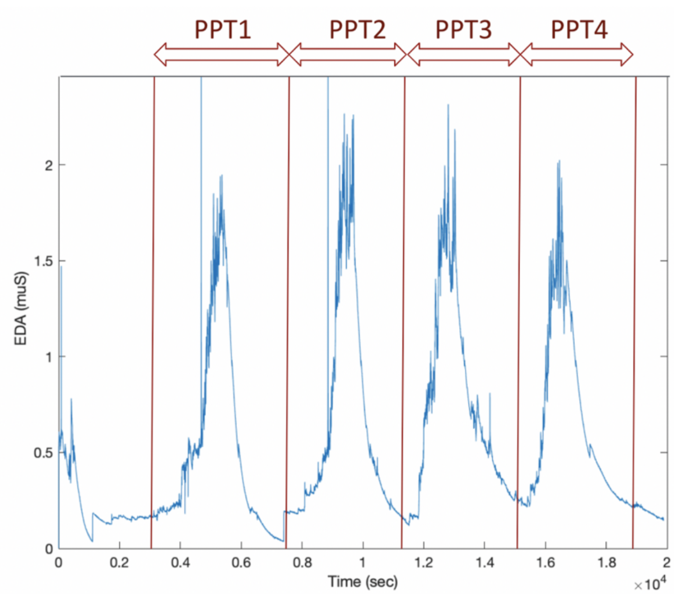
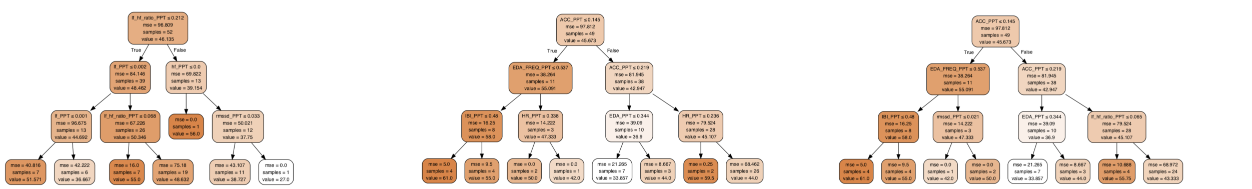


Figure 4.2: Example of electrodermal activity (EDA) during TEST treatments 1 to 4.

measures seem to outperform WWP measures. This would indicate that heart based measures and speech-based measures are better representatives of state-based PSA. It was also found that combining physiological and acoustic features appears to be more useful compared to including measures solely from a single modality. (Table 4.1) shows how the combination of features from WWP and CWP or from Acoustic and CWP provide highly significantly correlations.

In order to analyze why combining CWP and WWP based features increases the state-based anxiety prediction significantly, decision tree regressions were also conducted. Visual inspection of the decision trees (Table 4.3) indicated that the WWP based features i.e., the acceleration feature (root level node) and EDA based metrics (root child nodes) prove to be most informative and guide the regression. The heart rate feature from WWP is found to be less informative and is present at the second level of the tree. On combining the CWP features to the WWP features, this heart rate feature is replaced by the ECG based RMSSD and LF/HF ratio features from the CWP. This could suggest that the chest worn device provides more reliable estimate of the ECG and therefore of the sympathetic activity, thus integrating the CWP based measures provides more reliable state-based anxiety prediction.



(a) Actiwave based features (b) E4 based features (c) Actiwave & E4 based features

Figure 4.3: Decision trees created from physiological features extracted from the separate wearable modalities (a) CWP and (b) WWP as well as their combination (c) CWP & WWP.

4.0.2 Effect of individual-specific factors on PSA

4.0.2.1 Linear regression with interaction effects

Results suggest that PSA, as quantified by self-reports and bio-behavioral indices, is affected by a variety of individual and contextual factors. The interaction analysis results showcase few but significant associations between individual factors, physiology and state-based anxiety, for example, for individuals with high-trait anxiety, the parasympathetic activity (captured by RMSSD in heart beat occurrence) is not likely to get activated when they are engaged in a public speaking task as compared to individuals with high-trait anxiety (Fig. 4.4).

Table 4.1: Spearman's correlation between actual and estimated state-based anxiety using wrist-worn physiological (WWP), chest-worn physiological (CWP), and acoustic measures.

Bio-behavioural measures	**PRE session**	**POST session**
WWP	-0.05	-0.06
CWP	0.18	0.02
Acoustic	0.15	-0.37^{*}
WWP & CWP	0.32^{**}	0.25
WWP & Acoustic	0.14	-0.08
Acoustic & CWP	0.31^{**}	-0.19

*: p <0.05, **: p <0.01

4.0.2.2 Estimation of PSA from bio-behavioral measures augmented with individual and contextual factors

Results from regression experiments indicate that augmenting the original bio-behavioral features with individual and contextual factors benefits the estimation of state-based PSA, as shown in (Table 4.2a). Individual factors related to general trait-based anxiety and personality when combined with CWP features, increase the accuracy of PSA estimation from 0.18 (Table 4.1) to 0.36 (p <0.01) (Table 4.2a) during the PRE. Augmenting the WWP and CWP features with individual factors increased Spearman's correlation from 0.32 (p <0.01) (Table 4.1) to 0.38 (p <0.01) (Table 4.2a). Similarly, augmenting CWP and acoustic features with contextual factors such as age, education level etc. increases their Spearman's correlation from 0.31 (p <0.01) (Table 4.1) to 0.62 (p <0.01) (Table 4.2b). A comparison of the these aforementioned significant correlation coefficients was also conducted using Fisher r-to-z transformation. Results showcased significant differences between certain correlation coefficients such as Acoustic & CWP vs. Acoustic & CWP augmented with contextual factors ($z = -1.78, p = 0.03$). Significant increase was also found for the POST, benefiting most of the models which were previously relying solely on the bio-behavioral features. Similar benefits were provided by augmenting the bio-behavioral feature space with contextual factors (Table 4.2b). Notably, combining acoustic features with informa-

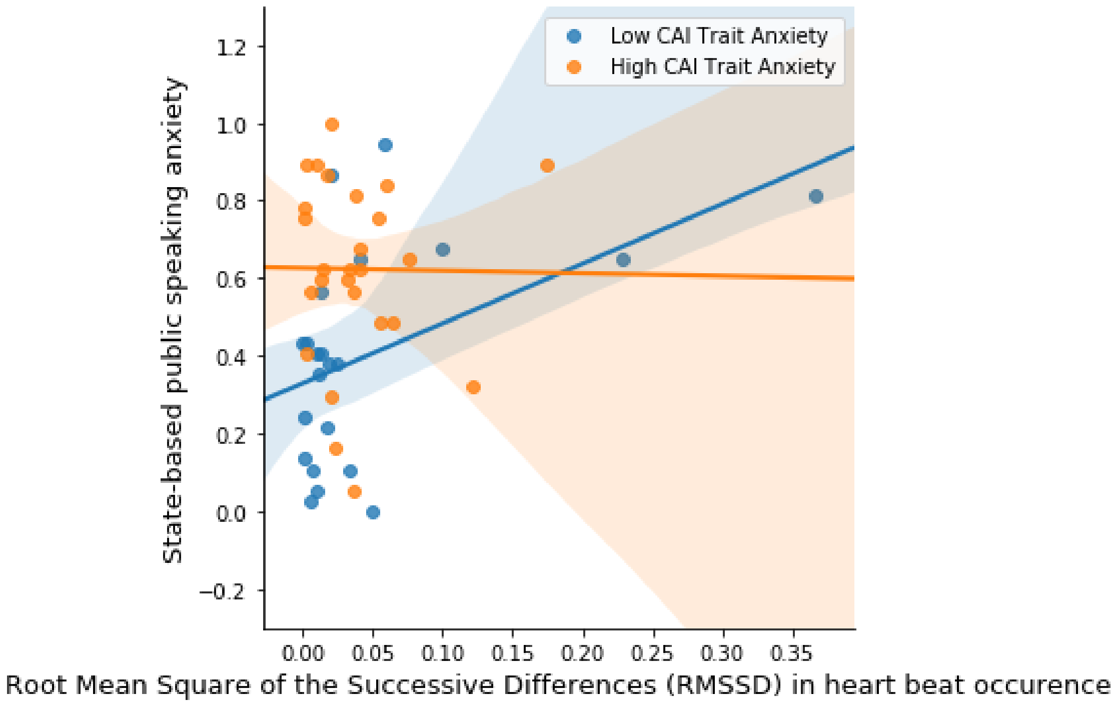
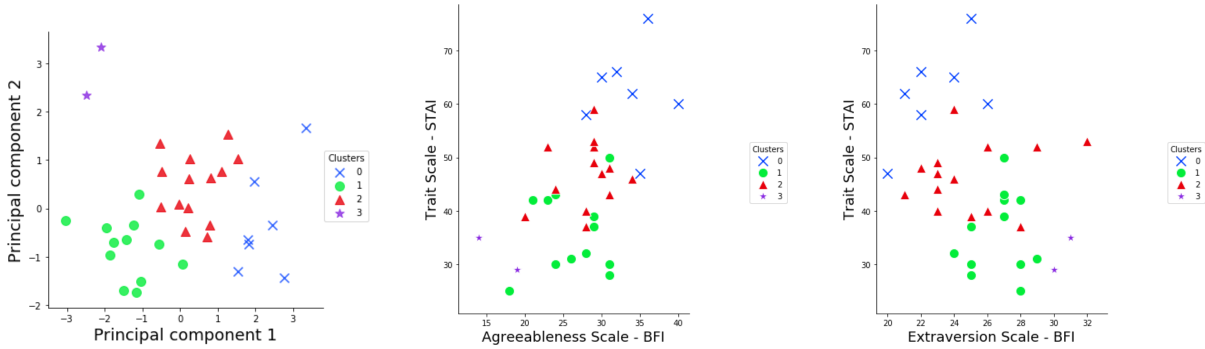


Figure 4.4: Predicted values of PRE session's state-based anxiety using the interaction based linear regression model for low/high score of trait-based Communication Anxiety Inventory (CAI) and physiology (RMSSD) measured via chest worn device.

tion on age, gender, native language, degree currently pursued, and recency of public speaking increases Spearman's correlation from 0.15 (Table 4.1) to 0.57 (p <0.01) (Table 4.2b) in the PRE. Contextual factors did not benefit results on the POST, potentially due to the fact that the small number of data samples in the POST might undermine the robustness of our results.

4.0.2.3 Group-specific clustering

For determining the sub-populations, K-means clustering was utilized. Results from K-Means clustering suggests the presence of various groups of participants. Fig. 4.5a depicts four distinctly separable clusters based on all individual factors. Visual inspection of the resulting clusters indicates groups of participants with high trait anxiety and high agreeableness (BFI peronality trait) (Fig. 4.5a), as well as high trait anxiety and low extraversion (BFI peronality trait) (Fig. 4.5b). Clustering based on contextual factors provided similar plots.



(a) PC1 vs. PC2 (b) Trait anxiety vs. Agreeableness (c) Trait anxiety vs. Extraversion

Figure 4.5: (a) Groups of participants as determined by K-Means clustering performed on the first two principal component dimensions of the individual factors. (b)-(c) Pairwise plots of individual factors: trait anxiety and Big Five Inventory (BFI) personality metric agreeableness/extraversion with grouping governed by K-means clustering on principal component dimensions of all the individual factors.

4.0.2.4 Identifying PSA differences between groups of participants

Significant differences can be observed among participant groups based on the various individual and contextual factors with respect to their PSA (Table 4.3). Participants who had given a presentation 4–8 times in the last 3 months, reported significantly higher PPP scores (mean= 17.80, stand. dev= 3.35) compared to participants who had presented only 1–3 times during the same duration (mean= 12.85, std= 3.90). Undergraduate students reported significantly higher trait anxiety (CAI trait) (mean= 15.33, stand. dev= 3.53) and depict higher SCR frequency (mean= 13.02, stand. dev= 3.77) compared to graduate students (CAI trait; mean= 10.80, stand. dev= 2.28) (EDA frequency; mean= 6.10, stand. dev= 4.23). Participants of Asian ethnicity depicted increased shimmer in their speech (mean= 0.13, stand. dev= 0.03) compared to White/Caucasian participants (mean= 0.12, stand. dev= 0.02), potentially due to general phonological differences between the two groups. Hispanic/Latino participants reported significantly higher BSQ (mean= 2.15, stand. dev= 0.70) compared to White/Caucasian participants (mean= 1.78, stand. dev= 0.55), as well as a higher SAE (mean= 60.77, stand. dev= 8.92) compared to African American participants (mean= 47.83, stand. dev= 7.25). Male participants depicted higher shimmer

Table 4.2: Spearman's correlation between actual and predicted state-based anxiety based on linear regression per modality and their combination with individual/contextual factors.

(a) Bio-behavioral measures augmented with individual factors related to trait-based (T) and personality (P) scores

Bio-behavioural measures	Individual factors	PRE session	POST session
WWP	T,P	-0.77^{**}	0.54^{**}
CWP	T,P	0.36^{**}	0.60^{**}
Acoustic	T,P	0.11	0.51^{**}
WWP & CWP	T,P	0.38^{**}	0.50^{**}
WWP & Acoustic	T,P	0.01	0.35
Acoustic & CWP	T,P	0.22	0.38^{*}

(b) Bio-behavioral measures augmented with contextual factors related to age (A), gender (G), native language (L), ethnicity (E), recency of public speaking presentation (R), highest degree earned (H), and degree currently pursuing (D)

Bio-behavioural measures	Contextual factors	PRE session	POST session
WWP	H	0.47^{**}	-0.20
CWP	A,R	0.49^{**}	0.24
Acoustic	A,G,L,D,R	0.57^{**}	-0.35
WWP & CWP	A,H,R	0.36^{*}	-0.22
WWP & Acoustic	L,E,H,R	0.36^{*}	-0.54^{**}
Acoustic & CWP	L,E,A,R	0.62^{**}	-0.33

*: p <0.05, **: p <0.01

(mean= 0.14, stand. dev= 0.03) compared to females (mean= 0.12, stand. dev= 0.02).

4.0.2.5 Group-specific PSA models

Results obtained from the group-specific FNNs provide better performance compared to the general baseline FNN, as depicted in (Table 4.4), which supports the hypothesis that a population specific model constructed via individual factors outperforms a general model solely based on bio-behavioural indices. While FNNs refined based on individual-specific clustering marginally improve the Spearman's correlation, FNNs refined using individual and contextual-specific clus-

Table 4.3: ANOVA & T-test results for measuring significant differences in public speaking anxiety (PSA) between various groups of individuals with respect to self-reports and bio-behavioural indices.

a) ANOVA and T-test results for self-assessments

	Communication Anxiety Inventory (CAI) trait (Dyadic)	Brief fear of Negative Evaluation (BFNE)	Retience Wilingness to Communicate (RWTC)	Communication Anxiety Inventory (CAI) state	Body sensations questionnaire (BSQ)	Post Presentation Performance (PPP)	State Anxiety Enthusiasm (SAE)
Age	$f(2,50) = 0.80$	$f(2,50) = 0.07$	$f(2,50) = 0.34$	$f(2,50) = 1.57$	$f(2,50) = 1.31$	$f(2,50) = 0.18$	$f(2,50) = 1.77$
Gender	$t(50) = 0.49$	$t(50) = -1.23$	$t(50) = -1.61$	$t(50) = -0.42$	$t(50) = -0.57$	$t(50) = 0.24$	$t(50) = -1.17$
Ethnicity	$f(4,50) = 0.51$	$f(4,50) = 0.87$	$f(4,50) = 0.57$	$f(4,50) = 2.11$	$f(4,50)$=**3.49****	$f(4,50) = 0.61$	$f(4,50)$ =**3.27***
College	$f(8,50) = 0.35$	$f(8,50)$ =**2.16***	$f(8,50) = 1.09$	$f(8,50) = 1.25$	$f(8,50) = 0.35$	$f(8,50) = 0.57$	$f(8,50) = 1.01$
Native language	$f(3,50) = 1.85$	$f(3,50) = 0.42$	$f(3,50) = 0.85$	$f(3,50) = 0.26$	$f(3,50) = 0.34$	$f(3,50) = 2.07$	$f(3,50) = 0.22$
Highest education	$f(3,50)$ =**2.95***	$f(3,50) = 0.19$	$f(3,50) = 0.14$	$f(3,50) = 0.07$	$f(3,50) = 0.17$	$f(3,50) = 0.44$	$f(3,50) = 0.15$
Presentation in last 3 months	$f(3,50) = 1.35$	$f(3,50) = 2.04$	$f(3,50) = 0.99$	$f(3,50) = 1.91$	$f(3,50) = 1.18$	$f(3,50)$ =**8.04****	$f(3,50)$ =**4.22****

* $p < 0.05$. ** $p < 0.01$

b) ANOVA and T-test results for bio-behavioural indices.

	Body temperature	Skin conductance response (SCR) frequency	Root mean square of successive differences (RMSSD) of R-R intervals	Speech jitter	Speech shimmer
Age	$f(2,50) = 1.72$	$f(2,50) = 1.77$	$f(2,50) = 0.44$	$f(2,50) = 0.16$	$f(2,26) = 0.26$
Gender	$t(50) = -0.048$	$t(50) = 1.18$	$t(50) = 0.97$	$t(50) = 1.45$	$t(50)$ =**2.20***
Ethnicity	$f(4,50)$ =**2.66***	$f(4,50) = 1.71$	$f(4,50) = 1.08$	$f(4,50)$ =**3.11***	$f(4,50) = 0.88$
Native language	$f(3,50) = 2.15$	$f(3,50)$ =**3.05***	$f(3,50) = 0.08$	$f(3,50) = 0.69$	$f(3,50) = 1.12$
Highest education	$f(3,50) = 1.51$	$f(3,50)$ =**4.74****	$f(3,50) = 2.06$	$f(3,50) = 0.53$	$f(2,26) = 0.32$
College	$f(8,50) = 1.25$	$f(8,50) = 1.41$	$f(8,50) = 0.40$	$f(8,50) = 0.49$	$f(8,50) = 0.92$
Presentation in last 3 months	$f(3,50) = 1.32$	$f(3,50) = 0.43$	$f(3,50)$ =**7.57****	$f(3,50) = 0.25$	$f(3,50) = 0.47$

* $p < 0.05$. ** $p < 0.01$

ters depict significant benefits, yielding a final Spearman's correlation of 0.55 (p <0.05) compared to 0.10 from the general FNNs. It is important to note that the sample size for the FNN decreased when combining the individual and contextual factors, because of missing data for some participants. This imbalance was taken into account and the context-based FNN was also tested on the reduced data set, which included the 18 participants whose individual and contextual metrics were both available, and results were found to be consistent with Table 4.4.

4.0.3 Examining effect of VR stimuli on PSA

4.0.3.1 Comparing PSA before and after the VR sessions

Significant differences with respect to self-reported and bio-behavioral indices were found before and after the 8 VR sessions (TEST treatment) (Table 4.5). The corresponding measures were obtained during public speaking presentations in front of a real audience, which occurred before (PRE treatment) and after (POST treatment) the VR sessions. Results suggest a significant re-

Table 4.4: Spearman's correlation between the actual and predicted state-based anxiety measures based on the group-specific feed-forward neural network (FNN) models.

Type of FNN model	**Sample size**	**PRE session**
Context-based FNN	35	-0.31
Baseline FNN	35	-0.39^{*}
individual FNN	21	0.15
Baseline FNN	21	0.02
Context & Trait-based FNN	18	0.55^{*}
Baseline FNN	18	0.10

*: p <0.05, **: p <0.01

duction in terms of self-reported state-based PSA (CAI, SAE) between the PRE and the POST, indicating that participants felt less stressed when presenting in front of the real audience after experiencing the VR sessions. It is also noteworthy that participants reported a reduction with approaching significance ($p = 0.06$) in terms of trait-based PSA, as obtained from the PRPSA metric, which might suggest a long-term usage of the proposed VR exposure. This decrease of 12.25% in trait-based PSA and 14% in state-based PSA via VR exposure is in agreement with previous studies which report that individuals who practice public speaking in VR based environments show a relative decrease of approx. 21% in their trait and 19% decrease in their state anxiety compared to a 4% decrease when practicing sans any VR support [39]. The presented results further reflect a significant reduction in SCR frequency and heart rate between the PRE and the POST, suggesting a reduction in the amount of sympathetic activity related to the fight-or-flight response during the POST. Although there are significant differences between the PRE and POST treatments with respect to the jitter and shimmer, the difference is not in the expected direction. Jitter, a measure related to the variations of fundamental frequency and speech breathiness [79], has increased during the POST compared to the PRE. This might be due to the fact that participants might have been more eager to touch upon as many discussion points as possible and show improved public speaking skills in front of the real audience during the POST, which might have caused the increased breathiness in their voice. While the limited number of samples (n =27) based on which this analysis is performed does not provide conclusive results, these findings indicate that systematic exposure to the public speaking through VR stimuli might be able to alleviate PSA.

Table 4.5: T-test results comparing significant differences between PRE and POST, before and after the virtual reality (VR) sessions, with respect to self-reported and bio-behavioral measures.

Self-reported measures	**PRE session**	**POST session**	**T-test results**
Communication Anxiety Inventory (CAI), State Scale	46.25	39.74	$t(26) = 2.33^{**}$
State-Anxiety Enthusiasm Scale (SAE)	55.66	48.14	$t(26) = 2.69^{**}$
Personal Report of Public Speaking Anxiety (PRPSA)	104.85	92.00	$t(26) = 1.88^{\dagger}$
Bio-behavioral measures	**PRE session**	**POST session**	**T-test results**
Skin conductance response frequency	11.83	6.84	$t(26) = 3.33^{**}$
Heart rate	89.23	82.46	$t(26) = 2.28^{*}$
Body temperature	32.71	31.80	$t(26) = 1.85^{\dagger}$
Jitter	0.02	0.04	$t(26) = -2.84^{**}$
Shimmer	0.12	0.15	$t(26) = -2.96^{**}$

†: p <0.1, *: p <0.05, **: p <0.01

4.0.3.2 Effect of VR environment on bio-behavioral signals

Correlation analysis through pearson correlation showcases few significant correlations/associations between the various bio-behavioral measures and the self-reported state-based anxiety scores, performance scores and VR immersiveness/sense scores as depicted in Table 4.6. For example, results showcase that participants who reported a high VR sense score i.e., felt more immersed and present in the VR environment, showcased a higher skin conductance response (Pearson's correlation= 0.12, $p < 0.05$). Similarly, participants who reported a high preparation and performance (PPP) score showcased more acceleration, i.e., they gestured more, which is a characteristic

Table 4.6: Pearson's correlation between bio-behavioral measures and state-based anxiety scores performance scores and virtual reality (VR) immersiveness/sense scores during TEST sessions.

Bio-behavioural measures	**PPP survey**	**SAE Score**	**VR Sense**
Acceleration	0.13^{*}	-0.03	-0.18^{**}
EDA frequency	-0.01	-0.20^{**}	0.12^{*}
Voice prob	0.11	-0.25^{**}	0.23^{**}
Fundamental frequency	0.11	-0.26^{**}	0.25^{**}

*: p <0.05, **: p <0.01

of a confident speaker.

4.0.3.3 Identifying how PSA is affected by various VR settings

The Presentation simulator software provides various VR environments, the 12 different VR settings utilized in the user study of this work are presented in Table 4.7. The effect of these 12 settings on the self-reported state-based anxiety scores scores was examined by calculating the mean of the scores across all participants. Results show that the state-based anxiety score increases to the maximum for VR environment (SAE score $= 54.956$) which was a highly negative meeting room setting, where the VR avatars were in close proximity to the participant and their negative expression would be clearly seen. This setting also had the VR avatars whispering to each other during the presentation. Participants had the lowest state-based anxiety score (SAE score $=$ 48.850) under VR environment 7, which was small neutral theatre. Analysis of variance (ANOVA) tests were performed to identify any confounding effects of VR settings on the participants bio-behavioural indices and self-reports. Even though majority of the ANOVA results did not indicate any significant differences between the various populations divided on the basis of VR settings, few of the results were approaching significance, for example, based on self-reports participants showcased some significant difference ($f(3, 27) = 2.112, p = 0.123$) in their state-based anxiety (SAE) when grouped in terms of the type of audience reaction they experienced i.e., negative, positive or neutral. This finding again is in accordance with previous studies which showcased

Table 4.7: Mean of state-based anxiety scores in the TEST sessions for all participants grouped based on 12 virtual reality (VR) environments provided by presentation simulator software.

Virtual Reality setting	**Room type**	**Audience reaction**	**Audience size**	**Percent female**	**SAE Score**
1	Executive meeting room	Neutral	12	50	49.391
2	Executive meeting room	Positive	12	50	49.571
3	Executive meeting room	Negative	12	50	54.956
4	Classroom	Neutral	25	50	51.800
5	Classroom	Positive	25	50	51.350
6	Classroom	Negative	25	50	51.500
7	Small theater	Neutral	90	50	48.850
8	Large hotel room	Neutral	54	50	54.280
9	Large hotel room	Negative	54	50	52.058
10	Executive meeting room	Neutral	12	30	49.650
11	Classroom	Neutral	25	30	49.185
12	Large hotel room	Neutral	54	30	50.000

that a negative audience is more anxiety provoking for an individual in a VR setting [34, 44]. The self-reported anxiety was also found to be reaching a significant difference based on the number of females present in the room ($t(27) = 1.709, p = 0.089$). Similarly, participants showcased some significant differences in their heart rate when they were grouped based on the type of room they performed in ($f(4, 27) = 2.253, p = 0.083$). This finding again is in agreement with previous studies which have found that the room/audience size plays a vital factor in PSA, studies have found that greater size audiences introduce more anxiety in an individual [10].

4.0.3.4 5-month follow-up

In order to evaluate the long-term effects of systematic exposure on trait-based PSA via VR-based public speaking stimuli, a 5-month follow-up survey was done with the participants. Results showcase decreased trait-based anxiety scores across participants' when comparing their PRE treatments self-assessments and their follow-up self-assessments (Table 4.8). Results from statistical analysis also show significant differences between the participants' PRE and follow-up trait-based anxiety levels (PRPSA) (p< 0.05) as well a decrease in their fear of audience evaluation

Table 4.8: Comparison of various state-based anxiety scores across participants between PRE treatments and 5-month follow-up survey.

State-based anxiety surveys	**Prior to VR exposure**	**5-month post VR**
Communication Anxiety Inventory (CAI)	47.47	44.94
Brief fear of Negative Evaluation (BFNE)	43.94^{*}	37.41^{*}
State Trait Anxiety Inventory (STAI)	45.93^{**}	41.58^{**}
Personal Report of Public Speaking Anxiety (PRPSA)	104.82	94.35

*: p <0.05, **: p <0.01

(BFNE) ($p < 0.01$).

5. DISCUSSION

The results reported in this work should be considered in light of the following limitations. Majority of the analysis in this work is based on the data obtained during the PRE and the POST treatments and only a preliminary investigation has been conducted with the data obtained from the VR treatments. These VR investigations have therefore not accounted for certain aspects such as, (i) the relevancy of the presentation topic provided to the participant in relation to the VR setting in which they are performing has not been considered, for example, does giving a presentation on a Shakespearean play in a corporate board room VR setting decreases the participant's VR immersiveness and consequently affects their PSA?, (ii) baselines or control groups were not utilized to asses the effect of the VR intervention in this work, and (iii) The design of the current user study prevents one from assessing the effect of certain audience specific factors on an individual's PSA, for example, does performing in front of negative VR audiences help participants more in mitigating their PSA? Following these factors it is understood that a more thorough investigation of the data from VR sessions would be required to understand the evolution of the participant's PSA while undergoing VR treatments. The current analysis also has not examined participants' visual cues, such as facial expressions or body gestures, which is an important channel for reflecting the degree of PSA, such an investigation would be a part of future work. In addition, current work has not yet accounted for the fact that the differences found between the PRE and the POST treatments might be attributed to the habituation effect arising from conducting the 10 public speaking tasks in a relatively small span of time (i.e., 2 weeks). A possible future direction of this research will include comparing desensitization through VR stimuli with other forms of interventions (e.g., desensitization with real audience, desensitization combined with cognitive restructuring feedback) in order to understand whether such differences will still be present. Data from a control group will be also collected as part of our future work. Cognitive restructuring feedback is a method which aims to modify an individual's negative perception of a threatening stimuli [9,80]. Previous work has used such interventions for treating PSA through a client's discussion with a therapist regarding ways to modify their negative perception of public speaking [41].

Wearable and mobile devices can afford us a unique solution for providing cognitive restructuring feedback in-the-moment, when it is needed the most. The hypothesis is that in-the-moment feedback would be able to change an individual's thought process, suppress their irrational fears, and direct them toward a healthier perception of public speaking. A future extension of this thesis would be to design a system where bio-behavioural indices and the population-specific PSA models are used to predict an individual's state-based PSA in real-time and provide them with in-the-moment feedback. The results of this work lay the foundation to move towards this direction by understanding the different bio-behavioral expressions of PSA among individuals and designing population-specific machine learning models capable of taking these factors into account.

6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

This research examined quantifiable estimators of PSA and the effect of VR in alleviating anxiety during public speaking. Statistical analysis indicates high inter-individual variability in the way participants perceive and experience PSA. Incorporating individual and contextual factors (e.g., trait-based anxiety, age , primary language, etc.) into machine learning models–either in the feature space, or through model adaptation–can improve PSA estimation. It was also identified that trait- and context-based factors combined together provide more predictive power to identify state-based anxiety compared to when used independently. Results demonstrate that systematic exposure to public speaking, implemented via VR, can help alleviate PSA in terms of self reports and bio-behavioral indices. In future work of this research, along with the individual-specific factors we will also integrate cognitive aptitude in the PSA models. Finally, we will obtain momentary PSA annotations from observational coding and design systems that can predict PSA in real-time, which will provide the foundation for in-the-moment PSA interventions.

6.2 Directions for future work

This work lays the foundation for adaptive individualized in-the-moment interventions systems for other types of social communication disorders using systematic exposure (e.g., through VR stimuli), relaxation methods, and cognitive restructuring feedback. The current design of population-specific computational models will contribute to the modeling of human behavior in everyday life through passively collected bio-behavioral indices and the data-augmented provision of in-the-moment feedback can result in behavioral change for empowering education and health. Examples of such applications include the monitoring and interventions for family well-being (e.g., conflict management between couples), pre-diabetic patients (e.g., modification of eating behaviors for glucose regulation), student learning (e.g., personalized classroom experience). The presented work will serve as a platform for achieving this goal, as it provides a foundation for computational models of human behavior and automated algorithms for personalized interventions in structured realistic settings with the potential to be extended in real-life.

REFERENCES

[1] F. R. Schneier, “Social anxiety disorder,” *New England Journal of Medicine*, vol. 355, no. 10, pp. 1029–1036, 2006.

[2] N. E. Betz, “Contributions of self-efficacy theory to career counseling: A personal perspective,” *The Career Development Quarterly*, vol. 52, no. 4, pp. 340–353, 2004.

[3] N. C. C. for Mental Health *et al.*, *Social anxiety disorder: the NICE guideline on recognition, assessment and treatment*. Royal College of Psychiatrists, 2013.

[4] C. Gallo, “New survey: 70% say presentation skills are critical for career success. forbes,” 2014.

[5] K. K. Dwyer and M. M. Davidson, “Is public speaking really more feared than death?,” *Communication Research Reports*, vol. 29, no. 2, pp. 99–107, 2012.

[6] K. Boonkit, “Enhancing the development of speaking skills for non-native speakers of english,” *Procedia-social and behavioral sciences*, vol. 2, no. 2, pp. 1305–1309, 2010.

[7] E. Carrillo, L. Moya-Albiol, E. González-Bono, A. Salvador, J. Ricarte, and J. Gómez-Amor, “Gender differences in cardiovascular and electrodermal responses to public speaking task: the role of anxiety and mood states,” *International Journal of Psychophysiology*, vol. 42, no. 3, pp. 253–264, 2001.

[8] R. Jones, *Communication in the real world: An introduction to communication studies*. The Saylor Foundation, 2013.

[9] M. T. Triscari, P. Faraci, D. Catalisano, V. D’Angelo, and V. Urso, “Effectiveness of cognitive behavioral therapy integrated with systematic desensitization, cognitive behavioral therapy combined with eye movement desensitization and reprocessing therapy, and cognitive behavioral therapy combined with virtual reality exposure therapy methods in the treatment of

flight anxiety: a randomized trial," *Neuropsychiatric disease and treatment*, vol. 11, p. 2591, 2015.

[10] G. D. Bodie, "A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety," *Communication education*, vol. 59, no. 1, pp. 70–105, 2010.

[11] R. C. Bednar, *Effect of public speaking anxiety on student academic learning in oral performance courses*. PhD thesis, 1991.

[12] J. A. Daly, A. L. Vangelisti, H. L. Neel, and P. D. Cavanaugh, "Pre-performance concerns associated with public speaking anxiety," *Communication Quarterly*, vol. 37, no. 1, pp. 39–53, 1989.

[13] M. B. Stein, J. R. Walker, and D. R. Forde, "Public-speaking fears in a community sample: Prevalence, impact on functioning, and diagnostic classification," *Archives of General Psychiatry*, vol. 53, no. 2, pp. 169–174, 1996.

[14] J. C. McCroskey, "The communication apprehension perspective," *Avoiding communication: Shyness, reticence, and communication apprehension*, pp. 13–38, 1984.

[15] M. L. Patterson and V. Ritts, "Social and communicative anxiety: A review and meta-analysis," *Annals of the International Communication Association*, vol. 20, no. 1, pp. 263–303, 1997.

[16] V. P. Richmond and M. Hickson, *Going public: A practical guide to public talk*. Allyn and Bacon, 2002.

[17] J. Andreassi, "Heart activity and behavior. ii. stress, emotions, motivation, personality, social factors, brain interactions, and conditioning," *Psychophysiology: Human behavior and physiological response (5 edn). Mahwal, NJ: Lawrence Erlbaum Associates, Inc., Publishers*, 2007.

[18] J. C. McCroskey, "Willingness to communicate, communication apprehension, and self-perceived communication competence: Conceptualizations and perspectives," *Avoiding com-*

munication: Shyness, reticence, and communication apprehension, vol. 2, pp. 191–216, 1997.

[19] M. Dickens and W. R. Parker, "An experimental study of certain physiological, introspective and rating-scale techniques for the measurement of stage fright," *Communications Monographs*, vol. 18, no. 4, pp. 251–259, 1951.

[20] R. R. Behnke and L. W. Carlile, "Heart rate as an index of speech anxiety," 1971.

[21] K. Clements and G. Turpin, "Physiological effects of public speaking assessed using a measure of palmar sweating.," *Journal of Psychophysiology*, 1996.

[22] L. E. Michelson and L. Ascher, *Anxiety and stress disorders: Cognitive-behavioral assessment and treatment.* Guilford Press, 1987.

[23] J. A. Daly, *Avoiding communication: Shyness, reticence, and communication apprehension*. Hampton Pr, 1997.

[24] A. Mulac and A. Sherman, "Conceptual foundations of the behavioral assessment of speech anxiety," *Western Journal of Speech Communication*, vol. 39, pp. 176–180, 1975.

[25] M. R. Lewin, D. W. McNeil, and J. M. Lipson, "Enduring without avoiding: Pauses and verbal dysfluencies in public speaking fear," *Journal of Psychopathology and Behavioral Assessment*, vol. 18, no. 4, pp. 387–402, 1996.

[26] L. Batrinca, G. Stratou, A. Shapiro, L.-P. Morency, and S. Scherer, "Cicero-towards a multimodal virtual audience platform for public speaking training," in *International workshop on intelligent virtual agents*, pp. 116–128, Springer, 2013.

[27] L. Chen, G. Feng, J. Joe, C. W. Leong, C. Kitchen, and C. M. Lee, "Towards automated assessment of public speaking skills using multimodal cues," in *Proc. ICMI*, pp. 200–203, ACM, 2014.

[28] M. J. Beatty and R. R. Behnke, “Effects of public speaking trait anxiety and intensity of speaking task on heart rate during performance,” *Human Communication Research*, vol. 18, no. 2, pp. 147–176, 1991.

[29] R. J. Croft, C. J. Gonsalvez, J. Gander, L. Lechem, and R. J. Barry, “Differential relations between heart rate and skin conductance, and public speaking anxiety,” *Journal of behavior therapy and experimental psychiatry*, vol. 35, no. 3, pp. 259–271, 2004.

[30] U. Dimberg, M. Fredrikson, and O. Lundquist, “Autonomic reactions to social and neutral stimuli in subjects high and low in public speaking fear,” *Biological Psychology*, vol. 23, no. 3, pp. 223–233, 1986.

[31] C. Kirschbaum, J. C. Prussner, A. A. Stone, I. Federenko, J. Gaab, D. Lintz, N. Schommer, and D. H. Hellhammer, “Persistent high cortisol responses to repeated psychological stress in a subpopulation of healthy men,” *Psychosomatic medicine*, vol. 57, no. 5, pp. 468–474, 1995.

[32] A. Schwerdtfeger, “Predicting autonomic reactivity to public speaking: don’t get fixed on self-report data!,” *International Journal of Psychophysiology*, vol. 52, no. 3, pp. 217–224, 2004.

[33] R. R. Behnke and C. R. Sawyer, “Anticipatory anxiety patterns for male and female public speakers,” *Communication education*, vol. 49, no. 2, pp. 187–195, 2000.

[34] D. P. Pertaub, M. Slater, and C. Barker, “An experiment on public speaking anxiety in response to three different types of virtual audience,” *Presence: Teleoperators & Virtual Environments*, vol. 11, no. 1, pp. 68–78, 2002.

[35] J. Ayres and T. Hopf, “Visualization: Reducing speech anxiety and enhancing performance,” *Communication Reports*, vol. 5, no. 1, pp. 1–10, 1992.

[36] G. Erdmann and S. Baumann, “Are psychophysiologic changes in the" public speaking" paradigm an expression of emotional stress?,” *Zeitschrift fur experimentelle Psychologie: Organ der Deutschen Gesellschaft fur Psychologie*, vol. 43, no. 2, pp. 224–255, 1996.

[37] A. W. Zuardi, J. A. d. S. Crippa, J. E. C. Hallak, and R. Gorayeb, "Human experimental anxiety: actual public speaking induces more intense physiological responses than simulated public speaking," *Brazilian Journal of Psychiatry*, vol. 35, no. 3, pp. 248–253, 2013.

[38] J. Ayres, T. Hopf, and D. M. Ayres, "An examination of whether imaging ability enhances the effectiveness of an intervention designed to reduce speech anxiety," *Communication Education*, vol. 43, no. 3, pp. 252–258, 1994.

[39] B. L. Heuett and K. B. Heuett, "Virtual reality therapy: A means of reducing public speaking anxiety," *International Journal of Humanities and Social Science*, vol. 1, no. 16, pp. 1–6, 2011.

[40] M. M. North, S. M. North, and J. R. Coble, "Virtual reality therapy: An effective treatment for psychological," *Virtual reality in neuro-psycho-physiology: Cognitive, clinical and methodological issues in assessment and rehabilitation*, vol. 44, p. 59, 1997.

[41] P. L. Anderson, E. Zimand, L. F. Hodges, and B. O. Rothbaum, "Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure," *Depression and anxiety*, vol. 22, no. 3, pp. 156–158, 2005.

[42] S. R. Harris, R. L. Kemmerling, and M. M. North, "Brief virtual reality therapy for public speaking anxiety," *Cyberpsychology & behavior*, vol. 5, no. 6, pp. 543–550, 2002.

[43] J. Diemer, A. Mühlberger, P. Pauli, and P. Zwanzger, "Virtual reality exposure in anxiety disorders: impact on psychophysiological reactivity," *The World Journal of Biological Psychiatry*, vol. 15, no. 6, pp. 427–442, 2014.

[44] M. Slater, D.-P. Pertaub, and A. Steed, "Public speaking in virtual reality: Facing an audience of avatars," *IEEE Computer Graphics and Applications*, vol. 19, no. 2, pp. 6–9, 1999.

[45] M. M. North, S. M. North, and J. R. Coble, "Virtual reality therapy: An effective treatment for psychological disorders.," 2002.

[46] M. M. North, S. M. North, and J. R. Coble, "Virtual reality therapy: An effective treatment for psychological," *Virtual reality in neuro-psycho-physiology: Cognitive, clinical and methodological issues in assessment and rehabilitation*, vol. 44, p. 59, 1997.

[47] D. Pertaub, M. Slater, and C. Barker, "An experiment on fear of public speaking in virtual reality," *Studies in health technology and informatics*, pp. 372–378, 2001.

[48] J. Diemer, A. Mühlberger, P. Pauli, and P. Zwanzger, "Virtual reality exposure in anxiety disorders: impact on psychophysiological reactivity," *The World Journal of Biological Psychiatry*, vol. 15, no. 6, pp. 427–442, 2014.

[49] M. Chollet, T. Wörtwein, L.-P. Morency, A. Shapiro, and S. Scherer, "Exploring feedback strategies to improve public speaking: An interactive virtual audience framework," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 1143–1154, ACM, 2015.

[50] M. Chollet, T. Wörtwein, L.-P. Morency, and S. Scherer, "A multimodal corpus for the assessment of public speaking ability and anxiety.," in *LREC*, 2016.

[51] "Empatica e4." `https://www.empatica.com/`, Accessed 1/20/2019.

[52] "Actiwave cardio monitor." `https://www.camntech.com/`, Accessed 1/20/2019.

[53] "Oculus rift." `https://www.oculus.com/?locale=en_US`, Accessed 1/20/2019.

[54] "Presentation simulator." `http://www.presentationsimulator.com/`, Accessed 1/20/2019.

[55] "Classroom background noise." `https://www.youtube.com/watch?v=ApYyTBsn2K0`, Accessed 1/20/2019.

[56] C. D. Spielberger, "State-trait anxiety inventory," *The Corsini encyclopedia of psychology*, pp. 1–1, 2010.

[57] S. Booth-Butterfield and M. Gould, "The communication anxiety inventory: Validation of state-and context-communication apprehension," *Communication Quarterly*, vol. 34, no. 2, pp. 194–205, 1986.

[58] J. C. McCroskey, "Measures of communication-bound anxiety," 1970.

[59] A. Tavoli, M. Melyani, M. Bakhtiari, G. H. Ghaedi, and A. Montazeri, "The brief fear of negative evaluation scale (BFNE): translation and validation study of the iranian version," *BMC psychiatry*, vol. 9, no. 1, p. 42, 2009.

[60] D. M. Clark, A. Wells, *et al.*, "A cognitive model of social phobia," *Social phobia: Diagnosis, assessment, and treatment*, vol. 41, no. 68, pp. 00022–3, 1995.

[61] R. M. Rapee and R. G. Heimberg, "A cognitive-behavioral model of anxiety in social phobia," *Behaviour research and therapy*, vol. 35, no. 8, pp. 741–756, 1997.

[62] D. Watson and R. Friend, "Measurement of social-evaluative anxiety.," *Journal of consulting and clinical psychology*, vol. 33, no. 4, p. 448, 1969.

[63] M. Yadav, M. N. Sakib, K. Feng, T. Chaspari, and A. Behzadan, "Virtual reality interfaces and population-specific models to mitigate public speaking anxiety," in *Proceedings of the 8th International Conference on Affective Computing Intelligent Interaction (ACII 2019)*, IEEE, 2019.

[64] M. Pörhölä, "Trait anxiety, experience, and the public speaking state responses of finnish university students," *Communication Research Reports*, vol. 14, no. 3, pp. 367–384, 1997.

[65] D. L. Chambless, G. C. Caputo, P. Bright, and R. Gallagher, "Assessment of fear of fear in phobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire.," *Journal of consulting and clinical psychology*, vol. 52, no. 6, p. 1090, 1984.

[66] O. P. John, S. Srivastava, *et al.*, "The big five trait taxonomy: History, measurement, and theoretical perspectives," *Handbook of personality: Theory and research*, vol. 2, no. 1999, pp. 102–138, 1999.

[67] J. Lessiter, J. Freeman, E. Keogh, and J. Davidoff, "A cross-media presence questionnaire: The ITC-Sense of Presence Inventory," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 282–297, 2001.

[68] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[69] M. Usoh, E. Catena, S. Arman, and M. Slater, "Using presence questionnaires in reality," *Presence: Teleoperators & Virtual Environments*, vol. 9, no. 5, pp. 497–503, 2000.

[70] "Fill outliers matlab function." `https://www.mathworks.com/help/matlab/ref/filloutliers.html`, Accessed 1/20/2019.

[71] "Biosppy bio signal processing in python." `https://biosppy.readthedocs.io/en/stable/`, Accessed 1/20/2019.

[72] "OpenSMILE audEERING." `https://www.audeering.com/opensmile/`, Accessed 1/20/2019.

[73] H. D. Critchley, "Electrodermal responses: what happens in the brain," *The Neuroscientist*, vol. 8, no. 2, pp. 132–142, 2002.

[74] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of neuroscience methods*, vol. 190, no. 1, pp. 80–91, 2010.

[75] "Python toolbox for heart rate variability." `https://pypi.org/project/pyhrv/`, Accessed 1/20/2019.

[76] R. Champseix, "Heart rate variability analysis." `https://pypi.org/project/hrv-analysis/`, 2018.

[77] S. Laborde, E. Mosley, and J. F. Thayer, "Heart rate variability and cardiac vagal tone in psychophysiological research–recommendations for experiment planning, data analysis, and data reporting," *Frontiers in psychology*, vol. 8, p. 213, 2017.

[78] G. E. Billman, “The lf/hf ratio does not accurately measure cardiac sympatho-vagal balance,” *Frontiers in physiology*, vol. 4, p. 26, 2013.

[79] J. Hillenbrand and R. A. Houde, “Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech,” *Journal of Speech, Language, and Hearing Research*, vol. 39, no. 2, pp. 311–321, 1996.

[80] D. A. Clark, “Cognitive restructuring,” *The Wiley handbook of cognitive behavioral therapy*, pp. 1–22, 2013.